

Dipartimento di Economia e Finanza
Cattedra di Experimental and Behavioural Economics

**IL RUOLO DELLA WORKING MEMORY A BREVE
TERMINE NELLA PERFORMANCE DEI COGNITIVE
REFLECTION TEST: UN ESPERIMENTO ECONOMICO**

Relatore

Prof.ssa Daniela Di Cagno

Candidato

Daniele Mitrani

Matricola

656051

Correlatore

Prof. Luca Panaccione

Anno Accademico 2014/2015

INDICE

ABSTRACT

1. INTRODUZIONE

2. LITERATURE REVIEW

3. EXPERIMENTAL DESIGN

3.1 Ipotesi

3.2 Partecipanti

3.3 Procedura

4. STATISTICHE DESCRITTIVE

4.1 Differenze di genere

5. RISULTATI

6. CONCLUSIONI

7. APPENDICE

BIBLIOGRAFIA E FONTI WEB

Introduzione

La Dual Process Theory (DPT) ha guadagnato crescente attenzione da parte degli economisti. Effettivamente, la DPT studia le decisioni economiche attraverso l'interazione tra due processi cognitivi e ci sono numerose eccellenti review su come le teorie dual-process operano nella vita sociale (Chaiken & Trope, 1999; Wegner & Bargh, 1998). In particolare, il principio centrale di questa teoria è che il comportamento e il processo decisionale sono determinati dall'interazione di un processo cognitivo euristico o Sistema 1 e uno deliberativo razionale o Sistema 2 (Stanovich & West, 2002; Kahneman, 2003). Hanno usato il termine "sistema" unicamente come etichetta per indicare una serie di processi cognitivi che possono essere distinti per la loro capacità di controllare l'attenzione nei contenuti economici in cui operano. Il Sistema 1 è caratterizzato da un processo decisionale senza il consumo di risorse cognitive, in quanto automatico, emotivo e impulsivo (Gilbert, 1999). E' in grado di verificarsi senza la necessità di un'elevata quantità di attenzione disponibile, e senza la consapevolezza dell'innescamento o il funzionamento di tale processo. Stanovich e West (2002) hanno sostenuto che l'esecuzione dei processi automatici è obbligatoria quando si incontrano determinati stimoli. Il processo cognitivo euristico ha come caratteristica correlata la fede nei pregiudizi ed è influenzato dai bias cognitivi (come l'impulsività). Dall'altra parte il Sistema 2 è caratterizzato da regole di riflessione razionale, consumando risorse cognitive (Gilbert, 1999). Questo processo si basa su funzioni esecutive, tra cui il mantenimento degli obiettivi e il monitoraggio dell'operato del Sistema 1. E' fondamentale per il recupero di conoscenze ed esperienze passate e tenute in memoria, e per il controllo dell'attenzione durante il processo decisionale, in quanto ha bisogno di un'elevata quantità di attenzione disponibile in un determinato momento. In particolare, nel modello dual-process assumiamo che il Sistema 1 propone rapidamente una risposta impulsiva al problema da risolvere e che il Sistema 2 controlla la qualità di queste proposte, le quali possono essere approvate, corrette o sostituite. La decisione che infine viene espressa viene chiamata impulsiva o euristica se i partecipanti conservano la proposta iniziale ipotizzata senza modificarla. Questa interpretazione riguardo i due sistemi è simile al "modello di correzione" proposto da Gilbert (1999). Il modello di correzione presuppone che un giudizio impulsivo è espresso apertamente solo se è sostenuto dal Sistema 2. In altre parole, si ammette il potenziale fallimento delle funzioni di monitoraggio del processo cognitivo deliberativo se la decisione espressa rimane ancorata sui pregiudizi iniziali sbagliati. Molto dipende dall'abilità

degli individui di controllare la loro attenzione. L'attenzione è controllata quando il Sistema 2 corregge efficacemente la risposta, cioè quando quella impulsiva è sbagliata.

La domanda che rimane è: quale test possiamo usare per coinvolgere entrambi i processi cognitivi?

Il Cognitive Reflection Test (CRT da qui in poi; Frederick, 2005) illustra proprio questa interazione. Si tratta di un test di natura quantitativa appositamente progettato per suscitare il processo cognitivo predominante nel ragionamento dei partecipanti.

CRT 1: Una rete e una palla costano complessivamente 110 \$. La rete costa 100 \$ in più della palla, quanto costa la palla? (risposta corretta: 5 \$)

CRT 2: Se 5 macchine producono 5 palloni in 5 ore, quanto tempo occorre a 100 macchine per produrre 100 palloni? (risposta corretta: 5 ore)

CRT 3: In un lago ci sono delle ninfee. Ogni giorno, la quantità di ninfee raddoppia. Se le ninfee ricoprono completamente il lago in 48 giorni, quanto tempo occorre alle ninfee per ricoprirne la metà? (risposta corretta: 47 giorni)

I CRT illustrano l'interazione tra i due processi cognitivi, in quanto non richiedono competenze specifiche di calcolo matematico, e si pongono in modo tale da indurre i partecipanti a fornire la risposta più impulsiva, anche se sbagliata. Precisamente, tra tutte le possibili risposte sbagliate che la gente possa dare, quelle impulsive (10, 100 e 24) dominano. Solo quelli con alta capacità di controllare l'attenzione scoprono la "trappola" innescando le funzioni di monitoraggio del processo cognitivo deliberativo e danno la risposta corretta. Per esempio tutti i ricercatori hanno misurato la performance nei tre CRT dando un punto per ogni risposta esatta. Ponti & Cueva et al. (2015) hanno misurato lo score dei CRT senza incentivi monetari e hanno trovato una media di 1,08 per i maschi e 0,55 per le femmine. Così come Bosch & Domenech et al. (2014), sempre senza incentivi, hanno trovato una media di 0,95 per i maschi e 0,58 per le femmine¹. Altri autori, invece, hanno misurato lo score dei CRT utilizzando gli incentivi monetari per premiare la risposta esatta. Di questi, Oechssler et al. (2009) trovarono una media di 2,2 per i maschi e 1,7 per le femmine. Sempre attraverso incentivi reali, Hoppe & Kusterer (2011) individuarono una media di 2,12 e 1,61 per i maschi e per le femmine rispettivamente. Nel nostro esperimento, usiamo incentivi monetari per premiare la risposta esatta e abbiamo

¹ Un ulteriore esempio l'hanno fornito Altmann, Falk & Grunewald (2015) trovando una media di 0,8 senza prendere in considerazione "l'effetto genere" e senza uno specifico pagamento per le performance nei CRT.

trovato una media di 1,27 per i maschi e 0,93 per le femmine, leggermente superiore alle medie senza incentivi reali. Quindi, abbiamo individuato un leggero miglioramento nelle performance attraverso i nostri incentivi monetari.

Un altro fattore cognitivo fondamentale è la Working Memory a breve termine, che non è distinta dalla memoria a breve termine. Quest'ultimo, è un vocabolo usato da Miller et al. (1956) per riferirsi alla memoria in quanto viene utilizzata per adottare un comportamento o prendere una decisione economica. La teoria di Miller prevede che la WM a breve termine umana abbia uno spazio di circa sette elementi (più o meno due). Lo spazio per la WM a breve termine è definito come la più lunga lista di elementi che una persona può ricordare nel corretto ordine dopo un breve periodo di tempo, anche in presenza di attività di distrazione e interferenza. Ricerche più recenti hanno mostrato che questo numero è approssimativamente accurato per gli studenti universitari, ma lo spazio della Working Memory a breve termine varia ampiamente con il tipo di popolazione testata, il materiale utilizzato (parole, numeri o immagini), il livello di carico cognitivo proposto e con il ricorso ad incentivi monetari reali (Cowan, 2001). In linea con Redick e Engle (2012), usiamo la percentuale di parole ricordate come proxy della Working Memory a breve termine. L'abilità di selezionare le informazioni (abilità nel controllo dell'attenzione) e di conservare queste informazioni di breve termine in uno stato accessibile (Working Memory a breve termine) sono aspetti critici delle nostre capacità cognitive (Barret, Tugade and Engle, 2004). Pertanto, assumiamo che la migliore Working Memory a breve termine è positivamente correlata con il maggiore numero di parole ricordate (attenzione). In altri termini, la nostra principale ipotesi congiunta è che il punteggio più alto di WM a breve termine permette un controllo maggiore dell'attenzione attivando le funzioni del processo cognitivo deliberativo (Sistema 2). Questa tesi è una delle prime che tenta di spiegare questi legami attraverso un esperimento economico, oltre ad essere il primo tentativo di mettere in relazione la WM a breve termine per le prestazioni cognitive coinvolgendo degli incentivi monetari. Nel 2004, Feldman, Engle e Tugade hanno sostenuto che una persona con un alto punteggio di WM a breve termine, dovrebbe avere le risorse sufficienti per focalizzare l'attenzione sulle informazioni obiettivo e per gestire la propria struttura di conoscenze di base pertinente. Al contrario, quelli più bassi in WM a breve termine possono essere meglio descritti dalla metafora degli "avidhi cognitivi" (Taylor, 1981). Gli avidhi cognitivi hanno risorse attentive fortemente limitate e, di conseguenza, adottano strategie che riducono la necessità di controllare l'attenzione. Anche se possono avere una serie di obiettivi o motivi, non hanno le risorse attentive sufficienti per mantenere l'elaborazione delle informazioni obiettivo rilevanti di fronte a situazioni complesse. L'autore suggerisce che le differenze comportamentali a livello

individuale catturate dalla DPT ispirano gli esperimenti economici e che esse possono, almeno in parte, essere spiegate attraverso differenze individuali nella WM a breve termine. Questa è solo una teoria che si è fondata sulle osservazioni degli errori cognitivi associati al processo cognitivo euristico. Gli esperimenti in laboratorio possono chiarire la nostra principale ipotesi, perché la modalità di elaborazione controllata delle informazioni disponibili deve essere definita in termini di punteggio nei CRT, piuttosto che in termini di esperienza personale di controllo o di volontà. La gente spesso pensa che sta controllando il proprio comportamento anche quando in realtà non è così (Wegner, 2002). Questa osservazione è in linea con l'emergente teoria in cui l'attenzione viene catturata automaticamente dagli stimoli soprattutto quando c'è una preparazione all'uso dell'attenzione che lo permette. Come risultato, il processo cognitivo deliberativo non può essere meramente l'inverso degli effetti dell'elaborazione impulsiva, ma può anche impedire (o consentire) l'espressione di giudizi che sono stati guidati dagli stimoli. L'implicazione è che non possiamo fare affidamento sulle domande di un questionario chiedendo ai partecipanti se avessero esercitato la volontà in modo controllato quando hanno effettuato la scelta. Per questo motivo, gli esperimenti in laboratorio possono meglio verificare la nostra ipotesi congiunta, in quanto consentono di osservare direttamente non solo l'esito della scelta (la risposta nei CRT), ma anche le caratteristiche aggiuntive del processo cognitivo, come la WM a breve termine.

Per verificare la nostra ipotesi principale, abbiamo impostato delle regressioni logit usando come variabile dipendente lo Score dei CRT il quale assume un valore pari a "1" se i partecipanti danno la risposta corretta e "0" altrimenti per ogni CRT. Questa variabile dummy stabilisce se i partecipanti hanno attivato efficacemente il processo cognitivo deliberativo dando la risposta giusta. Successivamente, abbiamo ampliato le analisi utilizzando un modo alternativo di guardare il dataset attraverso una nuova variabile dummy chiamata "CRT Score Impulsivo". La nuova variabile dipendente riflette se i partecipanti hanno usato in prevalenza il processo cognitivo euristico dando la risposta impulsiva. Per questa ragione, il CRT funge da "trappola cognitiva" causata dal Sistema 1 piuttosto che dal Sistema 2. In altre parole, lo score impulsivo dei CRT è inteso come misura dell'inabilità di sopprimere la risposta impulsiva erronea (Ponti & Cueva et al. 2015). Quindi, dovremmo ottenere un set di risultati opposti alla variabile dipendente precedente, in quanto le risposte impulsive manifestano la prevalenza del processo cognitivo euristico che conduce all'errore, con conseguente fallimento della funzione di monitoraggio del Sistema 2.

In aggiunta, analizziamo tre diversi effetti principali che possono influenzare i nostri risultati sulle prestazioni nella Working Memory a breve termine: l'Effetto Prova, l'Effetto Lunghezza della parola e l'Effetto Ordine. Il primo migliora le performance in quanto se i partecipanti hanno abbastanza tempo, possono ripetere silenziosamente in mente le parole per il successivo richiamo (Campbell & Baghaw, 2008). Intuitivamente, il secondo stabilisce che più la parola da ricordare è lunga e meno è la probabilità che hanno i partecipanti di ricordarla (Baddeley, Thomson & Buchanan, 1975). Il terzo si verifica quando ci sono numerosi elementi (come numeri, parole o immagini) da tenere in mente simultaneamente e tali elementi competono tra loro per il richiamo successivo. I nuovi elementi, di conseguenza, gradualmente “spingono” i precedenti fuori dalla memoria a meno che non sono attivamente protetti dalle interferenze o tramite “l'Effetto Prova” o tramite l'apprendimento focalizzando l'attenzione su di essi (Turner & Engle, 1989). Quindi, se l'ordine dei trattamenti è sempre lo stesso (prima 3 parole da ricordare, poi quattro e infine cinque), i partecipanti possono imparare dai primi trattamenti e di conseguenza ricordare più parole (Redick & Engle et al. 2012). E' ancora un tema aperto quello relativo all'interazione tra i bias cognitivi e gli incentivi monetari negli esperimenti. Questo problema è stato affrontato per la prima volta da Chamberlin (1948), egli suggerì un metodo per risolverlo, essenzialmente offrendo ai partecipanti un reale incentivo monetario. Sulla stessa linea, Smith nel 1976 ha enfatizzato l'importanza di offrire sufficienti incentivi monetari, al fine di evitare che i partecipanti si sottraggano ai “costi decisionali”, come il consumo di risorse cognitive durante la riflessione prolungata. Gli economisti in genere danno per scontato che il processo decisionale sia motivato principalmente da incentivi monetari, e che le decisioni economiche siano regolate principalmente da interessi reali e razionalità. In questo contesto, la razionalità è intesa come l'utilizzo delle informazioni disponibili in modo logico e sistematico al fine di fare scelte ottimali (Kahneman, 2002), ovvero dare la risposta giusta nei CRT.

Infine nei nostri risultati, confermiamo la prima principale ipotesi ovvero che il nostro fattore chiave (WM a breve termine) è fondamentale per comprendere il cambiamento degli Score nei CRT. In particolare quando il CRT Score riflette le risposte corrette, la performance sui CRT aumenta con una migliore Working Memory a breve termine. Abbiamo confermato anche l'Effetto Genere sui CRT, il quale si traduce in un vantaggio per i maschi nelle performance. Per la quantità di tempo a disposizione per ripetere le parole, potremmo dire che probabilmente i partecipanti non abbiano abbastanza tempo, ma non possiamo affermarlo con sicurezza. Mentre l'Effetto Lunghezza viene preso in considerazione attraverso una variabile dummy (come Baddeley, Thomson & Buchanan, 1975; La Point & Engle, 1990; Campbell et al. 2008).

Invece, abbiamo il sospetto che l'Effetto Ordine influenzi i nostri risultati. La nostra tesi è composta dalla letteratura di riferimento e dalle motivazioni che ci hanno spinto ad analizzare un particolare aspetto delle teorie dual-process, seguite dal nostro design dell'esperimento, dalle statistiche descrittive e dai i nostri principali risultati con la relativa discussione finale.

Design dell'esperimento

Ipotesi

In questa tesi esaminiamo le seguenti quattro ipotesi:

1. Le Performance nei CRT migliorano con la Working Memory a breve termine misurata attraverso le parole da ricordare.
2. L'Effetto di Genere deteriora le performance delle femmine nei CRT.
3. Il numero di parole ricordate (come proxy della WM a breve termine) si riduce con l'aumentare del numero delle parole da mantenere in memoria.
4. Il numero di parole ricordate (come proxy della WM a breve termine) si riduce con l'aumentare della lunghezza delle parole da mantenere in memoria.

La nostra **Ipotesi 1** è stata analizzata nello studio di Barret, Tugade e Engle (2004). Rappresenta un'ipotesi congiunta in quanto la letteratura teorica ipotizza che la Working Memory a breve termine è correlata con l'attenzione controllata che guida i partecipanti verso l'utilizzo del processo cognitivo deliberativo il quale migliora le performance nei CRT. Rigettare questa ipotesi potrebbe significare il rigetto di uno qualsiasi (o più di uno) degli elementi della teoria, quali:

- (i) La WM a breve termine può non aumentare l'attenzione controllata.
- (ii) L'attenzione non guida i partecipanti verso lo stato cognitivo deliberativo.
- (iii) Lo score nei CRT non migliora quando si usa lo stato cognitivo deliberativo.

La nostra **Ipotesi 2** è stata analizzata dalla letteratura e da Frederick (2005), e si fonda sulla forte componente di “Effetto Genere” (Maschio, Femmina) sulle prestazioni del CRT. I maschi hanno un punteggio superiore rispetto alle femmine. Rigettare questa ipotesi significa che non c'è alcuna differenza nelle performance nei CRT tra maschi e femmine.

La nostra **Ipotesi 3** è stata analizzata nello studio di Turner e Engle (1989), e si fonda sul deterioramento della Working Memory a breve termine con l'aumentare del numero di parole da tenere in mente. Rigettare questa ipotesi significa che l'aumento del numero delle parole da ricordare non influenza la percentuale di parole ricordate.

La nostra **Ipotesi 4** è stata analizzata nello studio di Baddeley, Thomson e Buchanan (1975), e si fonda sul deterioramento della Working Memory a breve termine con l'aumentare della lunghezza delle parole da ricordare. Rigettare questa ipotesi significa che l'aumento della lunghezza delle parole da tenere in mente non influenza la percentuale di parole ricordate.

Partecipanti

I partecipanti al nostro esperimento sono 134 studenti della University of Central Florida negli USA, di età compresa tra 18 e 36 con una media di 22.

Procedura

L'esperimento è stato condotto nell'arco di diversi mesi tra il 2007 e il 2008 per un totale di 12 sessioni. E' composto da due Task da completare ordinatamente uno alla volta. Tutti gli input di ciascun partecipante sono stati registrati sul computer. Prima di iniziare, i partecipanti hanno ricevuto le adeguate istruzioni di ogni attività. I partecipanti non hanno potuto utilizzare altri PC, tranne quello fornito dagli sperimentatori. Una volta che tutti sono stati registrati nel programma, l'esperimento è stato avviato attraverso un monitor centrale nella stanza. Al termine dell'esperimento ogni partecipante è stato pagato in privato.

Task 1: Operation Span Task

In questa sezione, usiamo un totale di 36 attività divise per 9 trattamenti per misurare la Working Memory a breve termine. Ogni attività è stata proiettata sullo schermo del PC personale, dividendolo orizzontalmente in due parti parallele. Nella parte alta è stata presentata un'equazione matematica chiedendo di indicare se tale equazione fosse corretta oppure no (come attività di interferenza). Mostriamo un esempio di un'equazione matematica usata:

$5 * 5 - 10 = 15 ?$ Yes No

Nella parte bassa dello schermo viene esposta una parola per il successivo richiamo (come proxy della Working Memory a breve termine). La lunghezza di tale parola varia a seconda della sessione.

Abbiamo diviso i 9 trattamenti in tre gruppi, ognuno contenente tre trattamenti. Nel primo gruppo sono state mostrate ad ogni trattamento tre schermate come quella descritta precedentemente, e alla fine è stata presentata una griglia per inserire le parole esposte poco tempo prima (per un totale di 9). Successivamente, nell'altro gruppo di tre trattamenti le schermate mostrate ad ogni trattamento sono diventate quattro, ed anche in questo caso alla fine dell'ultimo trattamento è stata esposta la relativa griglia da compilare con le parole (per un totale di 12). Nell'ultimo gruppo ne sono state mostrate cinque ad ogni trattamento, con l'ultimo trattamento seguito dalla griglia da compilare con le parole (per un totale di 15). Dopo aver completato il totale di 36 equazioni matematiche e parole da ricordare questo Task volge al termine. I partecipanti sono stati pagati 25 centesimi per ogni equazione matematica corretta e 25 centesimi per ogni parola ricordata correttamente (o compitata correttamente). Hanno potuto guadagnare un massimo di 18 \$ nel Task 1.

Task 2: Cognitive Reflection Test e Domande Demografiche

Questa è la sezione finale del nostro esperimento. Dopo aver completato il Task 1, il computer mostra un'ultima schermata in cui i tre CRT sono posti in cima allo schermo seguiti in basso dalle domande demografiche. Il Task 2 è formato da 22 domande, le cui prime tre sono i problemi esposti nei CRT e i partecipanti sono valutati e pagati 1\$ per ogni risposta corretta (come proxy che stabilisce l'efficiente utilizzo del processo cognitivo deliberativo). Mentre i partecipanti non hanno ricevuto nessun pagamento per le altre 19 domande demografiche, ma presumiamo che abbiano risposto il più sinceramente possibile.

Risultati e Conclusione

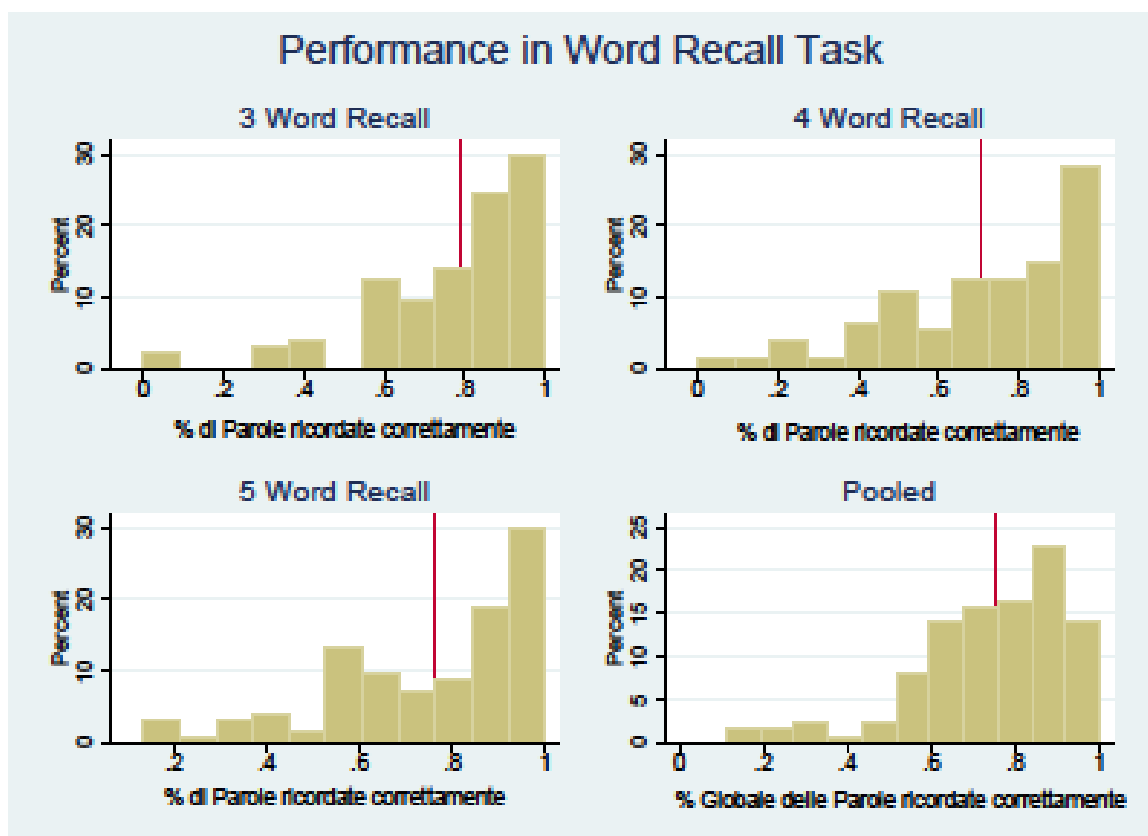
Iniziamo i nostri risultati generali verificando le caratteristiche della Working Memory a breve termine oggetto di analisi.

Verifichiamo la nostra **Ipotesi 4**:

Abbiamo osservato che la percentuale di parole ricordate si deteriora con l'aumentare della lunghezza delle parole tenute in mente. In quanto i partecipanti hanno ricordato una maggiore percentuale di parole "corte" rispetto a quelle "lunghe" (Baddeley, Thomson, e Buchanan, 1975).

Verifichiamo la nostra **Ipotesi 3**:

Grafico 1

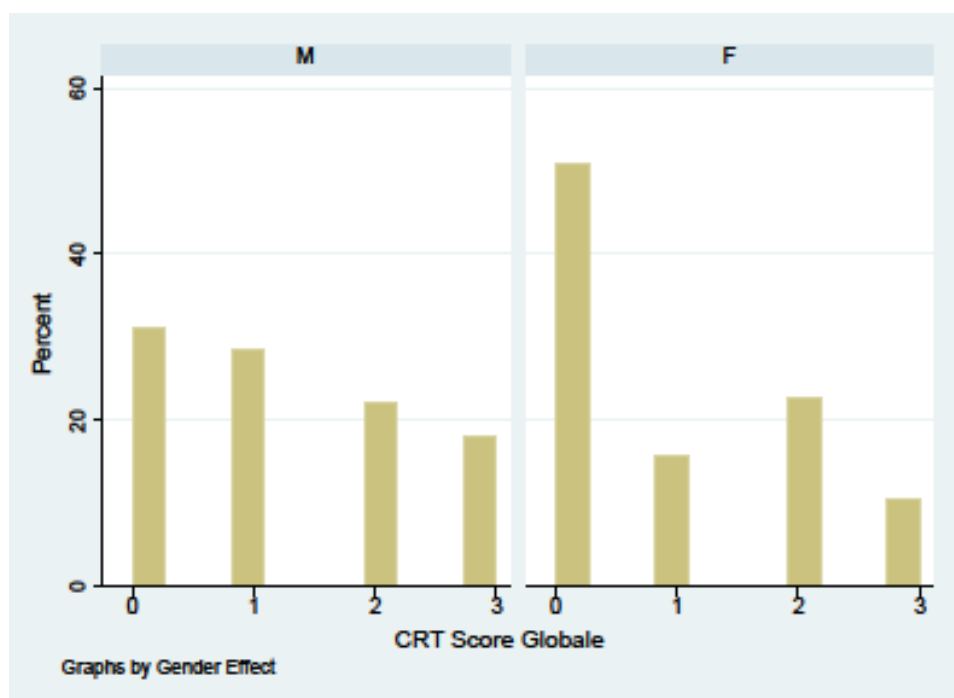


Non possiamo dimostrare l'ipotesi di Turner e Engle (1989), in cui la percentuale di parole ricordate si deteriora con l'aumentare del numero di parole tenute in mente. In particolare, l'Effetto Ordine per le equazioni aritmetiche è probabile che non si riferisca all'apprendimento, in quanto i partecipanti nel tempo ottengono performance minori. Questo suggerisce che i partecipanti abbiano più difficoltà nel risolvere tali problemi quando ce ne sono un numero maggiore, ma questa affermazione può essere confutata da un possibile Effetto Ordine (i partecipanti potrebbero annoiarsi o stancarsi). Per quanto riguarda la nostra proxy della WM a breve termine, invece, sembra che la performance su 4 parole sia minore confrontata con quella su 3 parole, mostrando un incremento di difficoltà. Il passaggio dei trattamenti da 4 a 5 parole sembra mostrare un incremento nelle performance, indicando apprendimento (Effetto Ordine). Per questa ragione, non possiamo confermare con certezza la nostra terza ipotesi.

Successivamente, analizziamo le performance sui Cognitive Reflection Test. Generiamo una variabile dipendente (CRT Score Globale) costruita assegnando un punto ai partecipanti per ogni risposta corretta sui tre CRT. E' la nostra proxy che stabilisce l'efficiente utilizzo del processo cognitivo deliberativo.

Verifichiamo la nostra **Ipotesi 2**:

Grafico 2 Performance on Global CRT Score by Gender Effect

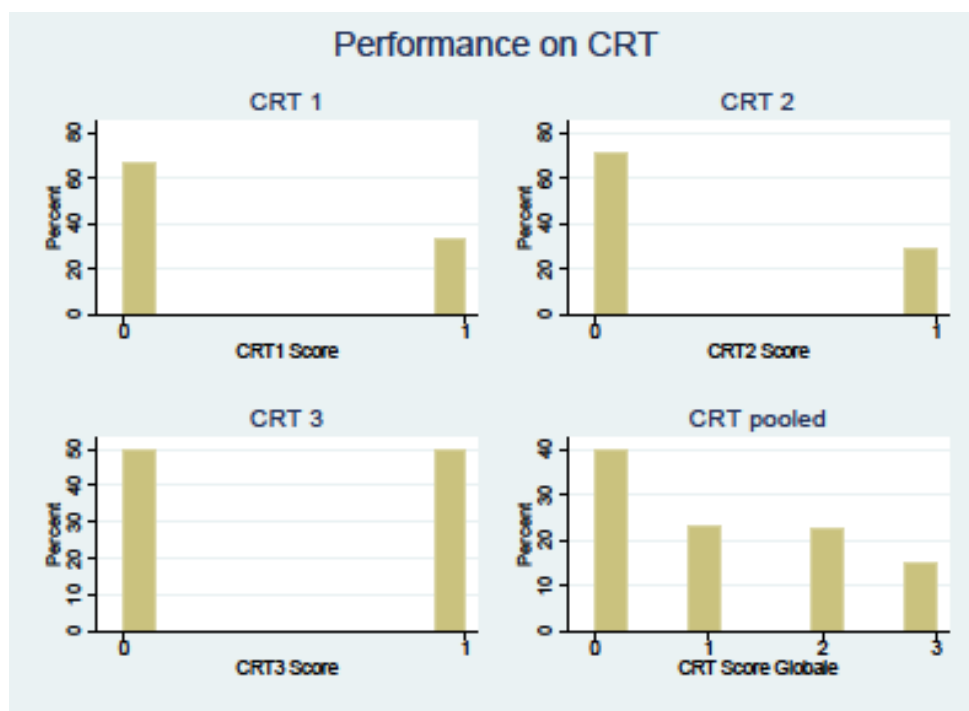


Il Grafico 2 mostra le performance nel CRT Score Globale, basato sulle risposte corrette, diviso per genere. Come si può notare, per esempio, la probabilità di zero risposte corrette è quasi il doppio per le femmine rispetto che per i maschi. Abbiamo trovato che l'Effetto Genere definito per la prima volta da Frederick (2005) nei CRT è solido. Quando il CRT Score riflette le risposte corrette, i maschi ottengono una performance particolarmente migliore rispetto alle femmine. Quando, invece, lo Score riflette le risposte impulsive, sono le femmine ad ottenere un punteggio più alto. Se un ricercatore è interessato ai risultati generali sui CRT, dovrebbe fare attenzione e dividere le prestazioni tra maschi e femmine. Per esempio, se si usa il criterio delle tre risposte impulsive, allora la performance è sproporzionatamente sbilanciata verso le femmine. Nel senso che, le femmine sembrano essere più impulsive, mentre i maschi probabilmente usano più spesso il processo cognitivo deliberativo quando prendono una decisione.

Verifichiamo la nostra **Ipotesi 1**:

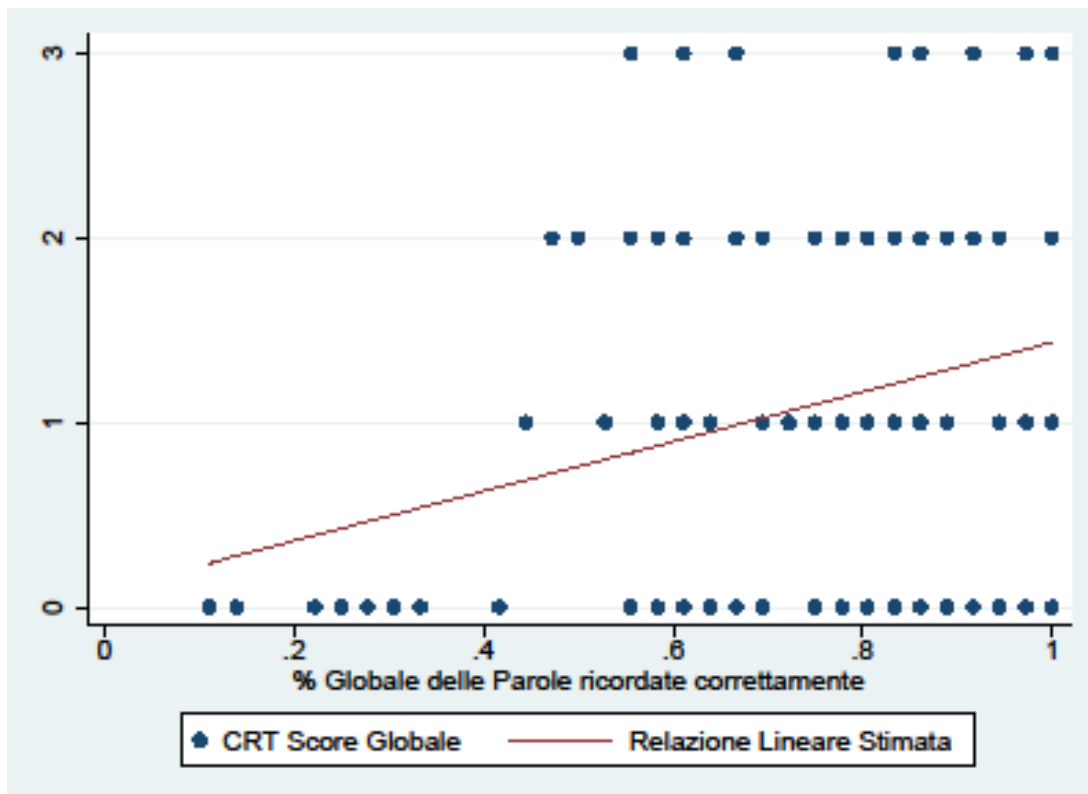
Nel Grafico 3, troviamo una media modesta di risposte corrette nei tre CRT, confermando la loro funzione di “trappola” cognitiva causata dalla prevalenza del Sistema 1 piuttosto che del Sistema 2.

Grafico 3



La nostra principale ipotesi congiunta è stata analizzata nello studio di Barret, Tugade e Engle (2004) e stabilisce che la Working Memory a breve termine (misurata dalla percentuale di parole correttamente ricordate) è correlata con l'attenzione controllata che guida i partecipanti verso l'utilizzo del processo cognitivo deliberativo il quale migliora le performance nei CRT.

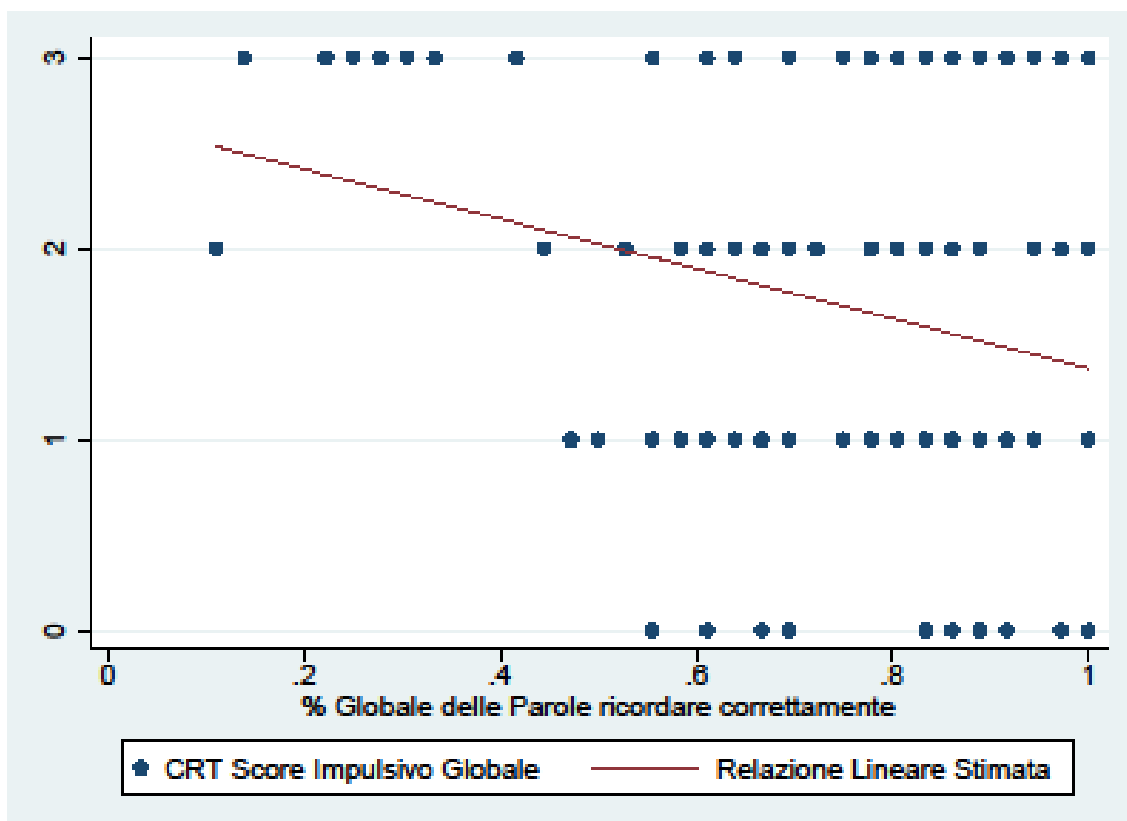
Grafico 3. Relazione tra il CRT Score Globale e la Working Memory a breve termine



Il Grafico 3 sembra dare un leggero supporto la nostra principale ipotesi ovvero che la maggiore percentuale di parole ricordate, come proxy della Working Memory a breve termine, migliora le performance sui CRT. Nel primo step, verifichiamo relazioni positive tra la percentuale di parole ricordate e l'educazione del padre e il reddito dei genitori, mentre tra la nostra proxy e l'educazione della madre evidenziamo una relazione negativa. Nel secondo step, regrediamo il CRT Score utilizzando come variabili indipendenti la percentuale di parole ricordate (come proxy della WM a breve termine), il genere, il GPA, il reddito dei genitori, l'educazione del padre e della madre, il major (che riflette le abilità matematiche dei partecipanti) e l'età. La nostra principale variabile indipendente è fondamentale per comprendere il cambiamento degli Score nei CRT. In particolare quando il CRT score riflette le risposte corrette, la performance

sui CRT aumenta con una migliore Working Memory a breve termine (misurata attraverso la percentuale di parole ricordate correttamente). Questo è coerente con la teoria di Barrett, Tugade e Engle (2004), in cui il punteggio più alto nella WM a breve termine permette un controllo maggiore dell'attenzione il quale attiva le funzioni di monitoraggio del processo cognitivo deliberativo (Sistema 2). Per completare i risultati nel secondo step, estendiamo l'analisi usando un modo alternativo di lavorare i dati attraverso una nuova variabile dipendente chiamata "CTR Score Impulsivo". Tra tutte le possibili risposte sbagliate che la gente possa dare, le risposte impulsive (10, 100 e 24) dominano. In altre parole, lo Score Impulsivo sui CRT è inteso come la misura dell'incapacità di sopprimere le risposte impulsive erronee (Ponti, Cueva et al. 2015). E' la nostra proxy che stabilisce il fallimento delle funzioni di monitoraggio del processo cognitivo deliberativo e la prevalenza del Sistema 1, in quanto costruita assegnando un punto ai partecipanti per ogni risposta impulsiva nei tre CRT.

Grafico 3 Relazione tra il CRT Score Impulsivo Globale e la WM a breve termine



I nostri risultati dimostrano che, in generale, la probabilità di dare la risposta corretta sembra più bassa della probabilità di dare la risposta impulsiva, confermando la teoria di Frederick (2005) e i risultati di Ponti e Cueva (2015). Il Grafico 3 mostra un ulteriore supporto alla nostra

ipotesi principale, osserviamo che la migliore WM a breve termine riduce la probabilità di dare la risposta impulsiva.

Il CRT è diventato sempre più popolare e il più usato nel predire la riflessione in esperimenti economici e in psicologia. Cosa rimane dopo la lettura di questa tesi è un ulteriore supporto alle teorie del dual-process. Nonostante le altre teorie economiche si focalizzino esclusivamente sul comportamento ottimale del pensiero razionale, la DPT conferma l'esistenza di due processi cognitivi nel pensiero razionale che competono tra loro per aiutarci a prendere le decisioni. Uno è un processo euristico, in cui non utilizziamo in modo ottimale le informazioni disponibili e siamo esposti ai bias cognitivi che limitano il nostro comportamento razionale. L'altro è un processo deliberativo, le informazioni sono analizzate efficientemente e il comportamento è completamente razionale. Il primo fattore chiave nella prevalenza dei due processi cognitivi è la Working Memory a breve termine che è positivamente correlata con la probabilità di usare il processo deliberativo per prendere decisioni. Il secondo fattore chiave è l'Effetto Genere, in cui le femmine hanno più probabilità di essere impulsive (bias cognitivo) e i maschi di usare la via completamente razionale. Tuttavia, i risultati osservati negli score dei CRT rimangono suscettibili di interpretazioni. La struttura causale può essere che il reddito e l'educazione dei genitori aumentino la capacità dei loro figli ad usare l'attenzione e a migliorare l'efficienza del processo cognitivo deliberativo quando necessario. Per questo motivo, queste variabili dovrebbero causare una Working Memory a breve termine migliore. Tuttavia, questa affermazione può invertire la causalità in modo tale da arrivare a conclusioni sbagliate. La causalità potrebbe essere che sia la WM a breve termine e sia altri fattori che dipendono dal reddito e dall'istruzione dei genitori concorrono a determinare le prestazioni sui CRT. Alcuni test per gli effetti demografici sono stati fatti spesso, ma una teoria per la quale andrebbero analizzati sembra mancare. In assenza di una tale teoria non è del tutto chiaro cosa esprime un effetto statisticamente non significativo. Sugeriamo che gli effetti demografici siano un altro fattore importante per comprendere meglio l'interazione tra la Working Memory a breve termine e le altre caratteristiche cognitive, emotive e culturali dell'individuo. Nell'immediato futuro ci proponiamo di continuare la stessa linea di ricerca incrementando la popolazione campionaria e la sua eterogeneità, al fine di ottenere una più ampia varietà di effetti demografici e di abilità cognitive.

Bibliografía e Fonti Web

- Altmann, S., Falk, A., & Grunewald, A. (2013). Incentives and Information as Driving Forces of Default Effects.
- Baddeley, A. D. (1986). Working memory. Oxford, England: Oxford University Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47–90). New York: Academic Press.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of verbal learning and verbal behavior*, 14(6), 575-589.
- Bar-Hillel, M. (1991). Commentary on Wolford, Taylor, and Beck: the conjunction fallacy?. *Memory & cognition*, 19(4), 412-414.
- Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological bulletin*, 130(4), 553.
- Bosch-Domènech, A., Brañas-Garza, P., & Espín, A. M. (2014). Can exposure to prenatal sex hormones (2D: 4D) predict cognitive reflection?. *Psychoneuroendocrinology*, 43, 1-10.
- Brañas-Garza, P., Kujal, P., & Lenkei, B. (2015). Cognitive Reflection Test: Whom, how, when.
- Campbell, R. D., & Bagshaw, M. (2008). *Human performance and limitations in aviation*. John Wiley & Sons.
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.
- Chamberlin, E. H. (1948). An experimental imperfect market. *The Journal of Political Economy*, 95-108.
- Conway, A. R. A., & Kane, M. J. (2001). Capacity, control and conflict: An individual differences perspective on attentional capture. In C. L. Folk, B. S. Gibson, & S. Bradley (Eds.), *Attraction, distraction and action: Multiple perspectives on attentional capture*. *Advances in psychology* (pp. 349–372). New York: Elsevier.
- Conway, A. R. A., Kane, M. J., Bunting, M.F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, 12, 769–786.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185.
- Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press
- Cueva Herrero, C., Iturbe-Ormaetxe Kortajarene, I., Mata-Pérez, E., Ponti, G., Sartarelli, M., & Yu, H. (2015). *Cognitive (Ir) reflection: New Experimental Evidence*. Instituto Valenciano de Investigaciones Económicas, SA (Ivie).
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, 11(1), 19-23.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol.44, pp.145–199). New York: Elsevier.
- Engle, R. W., Oransky, N. (1999). The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In R. Sternberg (Ed.), *The Nature of Cognition* (pp. 514-555). Cambridge, MA: MIT Press.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309-331.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGrawHill.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 25-42.
- Gilbert, D. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York: Guilford Press.
- Gilbert, D. (1991). How mental systems believe. *American Psychologist*, 46(2), 107–119.
- Gilbert, D. (1999). What the mind's not. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 3–11). New York: Guilford Press.
- Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin*, 118(2), 248.
- Holt, C. A., Porzio, M., & Song, M. Y. (2015). *Price Bubbles, Expectations, and Gender in Asset Markets: An Experiment*. University of Virginia working paper.
- Hoppe, E. I. and Kusterer, D. J. (2011). Behavioral Biases and Cognitive Reflection. *Economics Letters*, 110(2):97 – 100.
- James, W. (1890). 1950. *The principles of psychology*.
- Kahneman, D. (2000b). A psychological point of view: Violations of rational rules as a diagnostic of mental processes (commentary on Stanovich and West). *Behavioral and Brain Sciences*, 23, 681–683.

- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720.
- Kahneman, D., & Smith, V. (2002). Foundations of Behavioral and Experimental Economics. *Nobel Prize in Economics Documents, 1*.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133, 189–217.
- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working-memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp.21–48). New York: Oxford University Press.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133(2), 189.
- La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(6), 1118.
- Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis*. Psychology Press.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2), 81.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. Holt, Rinehart & Winston, New York.
- Oechssler, J., Roeder, A., & Schmitz, P. W. (2009). Cognitive Abilities and Behavioral Biases. *Journal of Economic Behavior & Organization*, 72(1):147–152.
- Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*.
- Savage, L. J. (1954). *The Foundations of Statistics*. New York: John Wiley
- Smith, C. A., & Morris, L. W. (1976). Effects of stimulative and sedative music on cognitive and emotional components of anxiety. *Psychological Reports*, 38(3c), 1187–1193.
- Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and brain sciences*, 23(05), 701-717.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and brain sciences*, 23(5), 645-726.
- Stanovich, K.E., & West, R.F. (2002). Individual differences in reasoning: Implications for the rationality debate? In T. Gilovich, D. W. Griffin, D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 421-440). New York: Cambridge University Press.
- Sternberg, R. J. (2000). The theory of successful intelligence. *Gifted education international*, 15(1), 4-21.
- Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 83–114). Hillsdale, NJ: Erlbaum.
- Toplak, M. E., West, R. F., and Stanovich, K. E. (2011). The Cognitive Reflection Test as a Predictor of Performance on Heuristics-and-biases Tasks. *Memory & Cognition*, 39(7):1275– 1289.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.
- Unsworth, N., & Engle, R.W. (2007). On the division of shortterm and working memory: An examination of simple and complex spans and their relation to higher-orderabilities. *Psychological Bulletin*, 133, 1038–1066.
- Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert & S. E. Fiske (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 446–496). New York: McGraw-Hill.
- Wittman, W. W. (1988). Multivariate reliability theory. In *Handbook of multivariate experimental psychology* (pp. 505-560). Springer US.

Dipartimento di Economia e Finanza
Cattedra di Experimental and Behavioural Economics

**IL RUOLO DELLA WORKING MEMORY A BREVE
TERMINE NELLA PERFORMANCE DEI COGNITIVE
REFLECTION TEST: UN ESPERIMENTO ECONOMICO**

Relatore

Prof.ssa Daniela Di Cagno

Candidato

Daniele Mitrani

Matricola

656051

Correlatore

Prof. Luca Panaccione

Anno Accademico 2014/2015

INDICE

ABSTRACT

1. INTRODUZIONE

2. LITERATURE REVIEW

3. DESIGN DELL'ESPERIMENTO

3.1 Ipotesi

3.2 Partecipanti

3.3 Procedura

4. STATISTICHE DESCRITTIVE

4.1 Differenze di genere

5. RISULTATI

6. CONCLUSIONI

7. APPENDICE

BIBLIOGRAFIA E FONTI WEB

Abstract

Gli economisti stanno prestando una maggiore attenzione alle teorie del processo duale della mente. Il principio centrale di queste teorie è che il comportamento e il processo decisionale sono determinati dall'interazione di un processo cognitivo euristico (Sistema 1) e uno deliberativo (Sistema 2). Il Cognitive Reflection Test (CRT) racchiude una combinazione di capacità cognitive, giudizi e processi decisionali e cattura importanti caratteristiche del pensiero razionale, le quali non sono misurate in altri test di intelligenza. Il nostro fattore chiave è la memoria di lavoro (Working Memory) a breve termine, la quale assumiamo sia positivamente correlata con il maggior numero di parole ricordate (attenzione). Pertanto, analizziamo se tale WM a breve termine permette un controllo maggiore dell'attenzione attivando le funzioni del processo cognitivo deliberativo (Sistema 2). Questa tesi è una delle prime che prova a spiegare questi legami attraverso un esperimento economico, oltre ad essere il primo tentativo di mettere in relazione la WM a breve termine per le prestazioni cognitive coinvolgendo degli incentivi monetari reali. Analizziamo alcuni degli effetti principali sulla WM a breve termine e sui CRT, i quali potrebbero influenzare la performance dei partecipanti.

Nei risultati, confermiamo la nostra principale ipotesi ovvero che il nostro fattore chiave è fondamentale per comprendere il cambiamento della performance nei CRT. In particolare, quando la performance riflette le risposte corrette, i risultati nei CRT aumentano con una migliore Working Memory a breve termine.

Introduzione

Gli economisti stanno prestando una maggiore attenzione alle teorie del processo duale della mente. Secondo queste teorie, esistono due modalità di elaborazione delle informazioni disponibili qualitativamente differenti che operano per formare giudizi, risolvere problemi o prendere decisioni. La prima è una modalità di elaborazione impulsiva, automatica, razionale benché influenzata dai bias cognitivi, in quanto non richiede un eccessivo uso dell'attenzione e di risorse cognitive e definisce il processo cognitivo euristico. La seconda è una modalità controllata, si basa su regole di riflessione razionale, in quanto richiede notevoli risorse attentive e cognitive come la riflessione prolungata e definisce il processo cognitivo deliberativo. Stanovich & West (2002) e Kahneman (2003) hanno utilizzato la Teoria del Processo Duale (DPT) per chiamare le due modalità: Sistema 1 e Sistema 2. Effettivamente, la DPT studia le decisioni economiche attraverso l'interazione tra i due sistemi e ci sono numerose eccellenti review su come le teorie dual-process operano nella vita sociale (Chaiken & Trope, 1999; Wegner & Bargh, 1998). Questa teoria è nata da una lunga tradizione letteraria in psicologia teorica con radici tracciate almeno fin dai suoi principi (1890, p. 451) dallo psicologo americano William James (1842-1910). In questa tesi si metterà alla prova un particolare aspetto delle teorie dual-process che sostiene un collegamento tra l'attenzione controllata e la memoria di lavoro (Working Memory) a breve termine, utilizzando i dati di laboratorio di un esperimento di economia.

Un altro fattore cognitivo fondamentale è la Working Memory a breve termine, che non è distinta dalla memoria a breve termine. Quest'ultimo, è un vocabolo usato da Miller et al. (1956) per riferirsi alla memoria in quanto viene utilizzata per adottare un comportamento o prendere una decisione economica. La teoria di Miller prevede che la WM a breve termine umana abbia uno spazio di circa sette elementi (più o meno due). Lo spazio per la WM a breve termine è definito come la più lunga lista di elementi che una persona può ricordare nel corretto ordine dopo un breve periodo di tempo, anche in presenza di attività di distrazione e interferenza. Ricerche più recenti hanno mostrato che questo numero è approssimativamente accurato per gli studenti universitari, ma lo spazio della Working Memory a breve termine varia ampiamente con il tipo di popolazione testata, il materiale utilizzato (parole, numeri o immagini), il livello di carico cognitivo proposto e con il ricorso ad incentivi monetari reali (Cowan, 2001). In linea con Redick e Engle (2012), usiamo la percentuale di parole ricordate come proxy della Working Memory a breve termine. L'abilità di selezionare le informazioni (abilità nel controllo

dell'attenzione) e di conservare queste informazioni di breve termine in uno stato accessibile (Working Memory a breve termine) sono aspetti critici delle nostre capacità cognitive (Barret, Tugade and Engle, 2004). Pertanto, assumiamo che la migliore Working Memory a breve termine è positivamente correlata con il maggiore numero di parole ricordate (attenzione). In altri termini, la nostra principale ipotesi congiunta è che il punteggio più alto di WM a breve termine permette un controllo maggiore dell'attenzione attivando le funzioni del processo cognitivo deliberativo (Sistema 2). Questa tesi è una delle prime che tenta di spiegare questi legami attraverso un esperimento economico, oltre ad essere il primo tentativo di mettere in relazione la WM a breve termine per le prestazioni cognitive coinvolgendo degli incentivi reali. Nel 2004, Feldman, Engle e Tugade hanno sostenuto che una persona con un alto punteggio di WM a breve termine, dovrebbe avere le risorse sufficienti per focalizzare l'attenzione sulle informazioni obiettivo e per gestire la propria struttura di conoscenze di base pertinente. Al contrario, quelli più bassi in WM a breve termine possono essere meglio descritti dalla metafora degli "avidhi cognitivi" (Taylor, 1981). Gli avidhi cognitivi hanno risorse attentive fortemente limitate e, di conseguenza, adottano strategie che riducono la necessità di controllare l'attenzione. Anche se possono avere una serie di obiettivi o motivi, non hanno le risorse attentive sufficienti per mantenere l'elaborazione delle informazioni obiettivo rilevanti di fronte a situazioni complesse. Gli autori suggeriscono che le differenze comportamentali a livello individuale catturate dalla DPT ispirano gli esperimenti economici e che esse possono, almeno in parte, essere spiegate attraverso differenze individuali nella WM a breve termine. Questa è solo una teoria che si è fondata sulle osservazioni degli errori cognitivi associati al processo cognitivo euristico. Secondo questo punto di vista, abbiamo basato la nostra ricerca sul "modello di correzione" proposto da Gilbert (1999), questo presuppone che un giudizio impulsivo è espresso apertamente solo se è sostenuto dal controllo del Sistema 2. Si suppone che il Sistema 1 e il Sistema 2 possano essere attivi contemporaneamente e che le operazioni cognitive automatiche e controllate competano per il controllo del processo decisionale. Se la capacità di controllare l'attenzione è bassa, e di conseguenza è minore la WM a breve termine, il Sistema 2 non dispone di risorse sufficienti per favorire il processo cognitivo deliberativo e monitorare il primo giudizio impulsivo.

Per valutare se le persone hanno risorse attentive sufficienti per attivare le funzioni di controllo del Sistema 2 usiamo il Cognitive Reflection Test (CRT seguito, Frederick, 2005). Il CRT illustra l'interazione tra i due processi cognitivi, in quanto non richiede competenze specifiche di calcolo matematico, e si pone in modo tale da indurre i partecipanti a fornire la risposta più impulsiva, anche se sbagliata. Precisamente, tra tutte le possibili risposte sbagliate che la gente

possa dare, quelle impulsive dominano. Solo quelli con alta capacità di controllare l'attenzione scoprono la "trappola" innescando le funzioni di monitoraggio del processo cognitivo deliberativo e danno la risposta giusta. Per esempio tutti i ricercatori hanno misurato la performance nei tre CRT dando un punto per ogni risposta esatta. Ponti & Cueva et al. (2015) hanno misurato lo score dei CRT senza incentivi monetari e hanno trovato una media di 1.08 per i maschi e 0.55 per le femmine. Così come Bosch & Domenech et al. (2014), sempre senza incentivi, hanno trovato una media di 0.95 per i maschi e 0.58 per le femmine². Altri autori, invece, hanno misurato lo score dei CRT utilizzando gli incentivi monetari per premiare la risposta esatta. Di questi, Oechssler et al. (2009) trovarono una media di 2.2 per i maschi e 1.7 per le femmine. Sempre attraverso incentivi reali, Hoppe & Kusterer (2011) individuarono una media di 2.12 e 1.61 per i maschi e per le femmine rispettivamente. Nel nostro esperimento, usiamo incentivi monetari per premiare la risposta esatta e abbiamo trovato una media di 1.27 per i maschi e 0.93 per le femmine, leggermente superiore alle medie senza incentivi reali. Quindi, abbiamo individuato un leggero miglioramento nelle performance attraverso i nostri incentivi monetari.

Gli esperimenti in laboratorio possono chiarire la nostra principale ipotesi, perché la modalità di elaborazione controllata delle informazioni disponibili deve essere definita in termini di punteggio nei CRT, piuttosto che in termini di esperienza personale di controllo o di volontà. La gente spesso pensa che sta controllando il proprio comportamento anche quando in realtà non è così (Wegner, 2002). Questa osservazione è in linea con l'emergente teoria in cui l'attenzione viene catturata automaticamente dagli stimoli soprattutto quando c'è una preparazione all'uso dell'attenzione che lo permette. Come risultato, il processo cognitivo deliberativo non può essere meramente l'inverso degli effetti dell'elaborazione impulsiva, ma può anche impedire (o consentire) l'espressione di giudizi che sono stati guidati dagli stimoli. L'implicazione è che non possiamo fare affidamento sulle domande di un questionario chiedendo ai partecipanti se avessero esercitato la volontà in modo controllato quando hanno effettuato la scelta. Per questo motivo, gli esperimenti in laboratorio possono meglio verificare la nostra ipotesi congiunta, in quanto consentono di osservare direttamente non solo l'esito della scelta (la risposta nei CRT), ma anche le caratteristiche aggiuntive del processo cognitivo, come la WM a breve termine.

² Un ulteriore esempio l'hanno fornito Altmann, Falk & Grunewald (2015) trovando una media di 0,8 senza prendere in considerazione "l'effetto genere" e senza uno specifico pagamento per le performance nei CRT.

E' molto più difficile valutare la Working Memory a breve termine. Engle e i suoi colleghi crearono e resero disponibili gratuitamente versioni computerizzate dei tre più usati metodi di misurazione della Working Memory a breve termine (Operation, Symmetry, and Reading Span Task), utilizzando diverse parole come elementi da ricordare. Essi fanno parte della famiglia dei Complex Span Task (CST) e tengono conto delle considerazioni psicologiche e teoriche conosciute che influenzano i punteggi su questi Task (Redick & Engle, 2012). I CST impegnano la Working Memory a breve termine, in quanto i partecipanti sono tenuti a mantenere le informazioni attive e rapidamente recuperabili, nonostante la loro attenzione sia periodicamente spostata verso le attività di interferenza. L'Operation Span Task consiste di due attività ripetute e mostrate contemporaneamente attraverso varie schermate di un PC. I partecipanti hanno osservato sia un'equazione aritmetica (come interferenza) da risolvere e sia una parola da ricordare e da richiamare poco tempo dopo. Una volta terminato un numero variabile di schermate, è stata presentata una griglia dove i partecipanti hanno inserito le parole nel corretto ordine in cui sono state presentate (come proxy della Working Memory a breve termine).

Per verificare la nostra ipotesi principale, abbiamo impostato delle regressioni logit usando come variabile dipendente lo Score dei CRT il quale assume un valore pari a "1" se i partecipanti danno la risposta corretta e "0" altrimenti per ogni CRT. Questa variabile dummy stabilisce se i partecipanti hanno attivato efficacemente il processo cognitivo deliberativo dando la risposta giusta. Successivamente, abbiamo ampliato le analisi utilizzando un modo alternativo di guardare il dataset attraverso una nuova variabile dummy chiamata "CRT Score Impulsivo". La nuova variabile dipendente riflette se i partecipanti hanno usato in prevalenza il processo cognitivo euristico dando la risposta impulsiva. Pertanto, per questa ragione il CRT funge da "trappola cognitiva" causata dal Sistema 1 piuttosto che dal Sistema 2. In altre parole, lo score impulsivo dei CRT è inteso come misura dell'inabilità di sopprimere la risposta impulsiva erronea (Ponti & Cueva et al. 2015). Quindi, dovremmo ottenere un set di risultati opposti alla variabile dipendente precedente, in quanto le risposte impulsive manifestano la prevalenza del processo cognitivo euristico che conduce all'errore, con conseguente fallimento della funzione di monitoraggio del Sistema 2.

In aggiunta, analizziamo tre diversi effetti principali che possono influenzare i nostri risultati sulle prestazioni nella Working Memory a breve termine: l'Effetto Prova, l'Effetto Lunghezza della parola e l'Effetto Ordine. Il primo migliora le performance in quanto se i partecipanti hanno abbastanza tempo, possono ripetere silenziosamente in mente le parole per il successivo richiamo (Campbell & Baghaw, 2008). Intuitivamente, il secondo stabilisce che più la parola

da ricordare è lunga e meno è la probabilità che hanno i partecipanti di ricordarla (Baddeley, Thomson & Buchanan, 1975). Il terzo si verifica quando ci sono numerosi elementi (come numeri, parole o immagini) da tenere in mente simultaneamente e tali elementi competono tra loro per il richiamo successivo. I nuovi elementi, di conseguenza, gradualmente “spingono” i precedenti fuori dalla memoria a meno che non sono attivamente protetti dalle interferenze o tramite “l’Effetto Prova” o tramite l’apprendimento focalizzando l’attenzione su di essi (Turner & Engle, 1989). Quindi, se l’ordine dei trattamenti è sempre lo stesso (prima 3 parole da ricordare, poi quattro e infine cinque), i partecipanti possono imparare dai primi trattamenti e di conseguenza ricordare più parole (Redick & Engle et al. 2012). E’ ancora un tema aperto quello relativo all’interazione tra i bias cognitivi e gli incentivi monetari negli esperimenti. Questo problema è stato affrontato per la prima volta da Chamberlin (1948), egli suggerì un metodo per risolverlo, essenzialmente offrendo ai partecipanti un reale incentivo monetario. Sulla stessa linea, Smith nel 1976 ha enfatizzato l’importanza di offrire sufficienti incentivi monetari, al fine di evitare che i partecipanti si sottraggano ai “costi decisionali”, come il consumo di risorse cognitive durante la riflessione prolungata. Gli economisti in genere danno per scontato che il processo decisionale sia motivato principalmente da incentivi monetari, e che le decisioni economiche siano regolate principalmente da interessi reali e razionalità. In questo contesto, la razionalità è intesa come l’utilizzo delle informazioni disponibili in modo logico e sistematico al fine di fare scelte ottimali (Kahneman, 2002), cioè dare la risposta giusta nei CRT.

Infine nei nostri risultati, confermiamo la prima principale ipotesi ovvero che il nostro fattore chiave (WM a breve termine) è fondamentale per comprendere il cambiamento degli Score nei CRT. In particolare quando il CRT Score riflette le risposte corrette, la performance sui CRT aumenta con una migliore Working Memory a breve termine. Abbiamo confermato anche l’Effetto Genere sui CRT, il quale si traduce in un vantaggio per i maschi nelle performance. Per la quantità di tempo a disposizione per ripetere le parole, potremmo dire che probabilmente i partecipanti non abbiano abbastanza tempo, ma non possiamo affermarlo con sicurezza. Mentre l’Effetto Lunghezza viene preso in considerazione attraverso una variabile dummy (come Baddeley, Thomson & Buchanan, 1975; La Point & Engle, 1990; Campbell et al. 2008). Invece, abbiamo il sospetto che l’Effetto Ordine influenzi i nostri risultati. La nostra tesi è composta dalla letteratura di riferimento e dalle motivazioni che ci hanno spinto ad analizzare un particolare aspetto delle teorie dual-process, seguite dal nostro design dell’esperimento, dalle statistiche descrittive e dai i nostri principali risultati con la relativa discussione finale.

Literature Review

La Dual Process Theory (DPT) ha guadagnato crescente attenzione da parte degli economisti. Il principio centrale di queste teorie è che il comportamento e il processo decisionale sono determinati dall'interazione di un processo cognitivo euristico o Sistema 1 e uno deliberativo razionale o Sistema 2 (Stanovich & West, 2002; Kahneman, 2003). Hanno usato il termine "sistema" unicamente come etichetta per indicare una serie di processi cognitivi che possono essere distinti per la loro capacità di controllare l'attenzione nei contenuti economici in cui operano. Il Sistema 1 è caratterizzato da un processo decisionale senza il consumo di risorse cognitive, in quanto automatico, emotivo e impulsivo (Gilbert, 1999). E' in grado di verificarsi senza la necessità di un'elevata quantità di attenzione disponibile, e senza la consapevolezza dell'inesco o il funzionamento di tale processo. Stanovich e West (2002) hanno sostenuto che l'esecuzione dei processi automatici è obbligatoria quando si incontrano determinati stimoli. Il processo cognitivo euristico ha come caratteristica correlata la fede nei pregiudizi ed è influenzato dai bias cognitivi. Dall'altra parte il Sistema 2 è caratterizzato da regole di riflessione razionale, consumando risorse cognitive (Gilbert, 1999). Questo processo si basa su funzioni esecutive, tra cui il mantenimento degli obiettivi e il monitoraggio dell'operato del Sistema 1. E' fondamentale per il recupero di conoscenze ed esperienze passate e tenute in memoria, e per il controllo dell'attenzione durante il processo decisionale, in quanto ha bisogno di un'elevata quantità di attenzione disponibile al momento. In particolare, nel modello dual-process assumiamo che il Sistema 1 propone rapidamente una risposta impulsiva al problema da risolvere e che il Sistema 2 controlla la qualità di queste proposte, le quali possono essere approvate, corrette o sostituite. La decisione che infine viene espressa viene chiamata impulsiva o euristica se i partecipanti conservano la proposta iniziale ipotizzata senza modificarla. Questa interpretazione riguardo i due sistemi è simile al "modello di correzione" proposto da Gilbert (1999). Il modello di correzione presuppone che un giudizio impulsivo è espresso apertamente solo se è sostenuto dal Sistema 2. Per esempio, il Test di Stroop illustra questa struttura a due sistemi. In questo test, degli osservatori sono stati istruiti a segnalare il colore in cui vengono stampate le parole. Sembra un compito facile in quanto non richiede una particolare quantità di attenzione disponibile o un controllo su essa. Gli osservatori tendono ad "inciampare" quando la parola in questione è il nome di un altro colore (ad esempio la parola "blu" stampata in verde). La difficoltà nasce dal fatto che la parola viene automaticamente letta e attiva una risposta impulsiva ("blu" in questo caso) che compete con la risposta corretta ("verde"). Gli errori sono

rari nel Test di Stroop, in quanto indice di un controllo e un monitoraggio generalmente efficiente del processo cognitivo razionale sulle risposte ai problemi di scelta, ma il conflitto genera ritardi e indecisioni. La soppressione delle risposte errate è faticosa e la sua efficacia può essere ridotta dallo stress e dalle interferenze (Gilbert, 1999). Quindi assumiamo che entrambi i sistemi si attivino e che i due processi cognitivi competano per il controllo della risposta da esprimere. In altre parole, si ammette il potenziale fallimento delle funzioni di monitoraggio del processo cognitivo razionale se la decisione espressa rimane ancorata sui pregiudizi iniziali. Molto dipende dall'abilità degli individui di controllare la loro attenzione. L'attenzione è controllata quando il Sistema 2 corregge efficacemente la risposta, cioè quando quella impulsiva è sbagliata.

La domanda che rimane è: quale test possiamo usare per coinvolgere entrambi i processi cognitivi?

Il Cognitive Reflection Test (CRT da qui in poi; Frederick, 2005) illustra proprio questa interazione. Si tratta di un test di natura quantitativa appositamente progettato per suscitare il processo cognitivo predominante nel ragionamento dei partecipanti.

CRT 1: Una rete e una palla costano complessivamente 110 \$. La rete costa 100 \$ in più della palla, quanto costa la palla? (risposta corretta: 5 \$)

CRT 2: Se 5 macchine producono 5 palloni in 5 ore, quanto tempo occorre a 100 macchine per produrre 100 palloni? (risposta corretta: 5 ore)

CRT 3: In un lago ci sono delle ninfee. Ogni giorno, la quantità di ninfee raddoppia. Se le ninfee ricoprono completamente il lago in 48 giorni, quanto tempo occorre alle ninfee per ricoprirne la metà? (risposta corretta: 47 giorni)

Questi test presentano una manipolazione cognitiva progettata per rendere inefficiente la funzione di monitoraggio del Sistema 2, inducendo il partecipante all'utilizzo del processo cognitivo euristico (a non riflettere abbastanza). La principale caratteristica del test è che, ad ogni domanda, è associata una risposta immediata ed "impulsiva" (10, 100 e 24 rispettivamente). Tali risposte impulsive e incorrette, possono essere selezionate da quei partecipanti che non hanno riflettuto abbastanza o non hanno controllato la loro attenzione. Frederick (2005) ha definito l'oggetto dei tre CRT come "facili" nel senso che le loro soluzioni sono di semplice comprensione una volta spiegate. Tuttavia, per raggiungere le risposte corrette è spesso richiesta la soppressione della risposta sbagliata che "impulsivamente" è stata richiamata dalla mente. Infatti in questo caso, trovare l'errore nella risposta impulsiva equivale

a risolvere il problema, per esempio nel CRT 1 quasi tutti quelli che non rispondono 10 \$ danno, dunque, la risposta esatta (5 \$). Pertanto, i CRT fungono da trappola cognitiva causata dall'utilizzo del processo cognitivo euristico piuttosto che da quello razionale. Questa affermazione è sostenuta da diversi elementi trovati ed esaminati da Frederick (2005). In primo luogo, tra tutte le possibili risposte sbagliate che le persone possano dare, quelle impulsive (10, 100 e 24) dominano. In secondo luogo, anche tra quelle che hanno risposto in modo corretto, la risposta sbagliata è stata spesso considerata prima, come è risultato dall'introspezione, rapporti verbali e scarabocchi sul margine del foglio (per esempio, 10 \$ è stato spesso trovato prima della scritta 5 \$, e mai il contrario). In terzo luogo, quando gli viene chiesto di giudicare la difficoltà del problema (stimando la percentuale di partecipanti che lo abbia correttamente risolto), i soggetti che hanno risposto impulsivamente hanno pensato che fosse il problema fosse molto facile. Per esempio, coloro che hanno risposto 10 \$ nel CRT 1 stimano che il 92 % delle persone ha risposto in modo corretto, mentre coloro che hanno dato la risposta giusta stimano che "solo" il 62 % ha fatto lo stesso (entrambe sono state sovrastime notevoli rispetto alle medie reali). Presumibilmente, questi ultimi hanno mentalmente barrato la risposta impulsiva e sapevano che non tutti si sarebbero comportati in questo modo, mentre chi ha dato la risposta impulsiva ha pensato che il problema era troppo facile da sbagliare (Frederick, 2005). In quarto luogo, le persone hanno maggiore successo nel rispondere a problemi analoghi che rappresentano un calcolo più complicato invitando una riflessione maggiore. Ad esempio, i partecipanti sbagliano molto più spesso il CRT 1 di quanto non sbagliano il test chiamato "della penna e della matita": "Una penna e una matita costano complessivamente 37 centesimi. La penna costa 13 centesimi in più rispetto alla matita. Quanto costa la penna? (con qualche calcolo si ottiene la risposta esatta "25 centesimi"). I dati di Frederick (2005) mostrano che ben il 78% rispondono positivamente contro circa il 31% del CRT 1. I partecipanti allo studio di Frederick (2005) a cui sono stati sottoposti i CRT erano 3428. L'esperimento includeva un incentivo monetario e comprendeva sia i CRT sia altri processi decisionali, quali le scelte in condizione di rischio intertemporale. Nelle pagine in cui apparivano i CRT, ai partecipanti è stato detto solo di rispondere correttamente al maggior numero di problemi, che presentavano diversi livelli di difficoltà. Molti dei partecipanti erano studenti universitari provenienti dal Nord America, questi hanno posto le basi per la maggior parte dei confronti statistici, e si è notato come le loro capacità cognitive erano molto più simili rispetto a quando l'esperimento era stato costituito sulla popolazione generale. Tuttavia, lo studio di Frederick è stato basato sull'analisi dei risultati degli studenti universitari che risultarono sorprendenti in quanto la media di risposta corretta su tre CRT è stata molto bassa (1.24). Tuttavia i suoi risultati potevano sottovalutare le

reali differenze nelle prestazioni usando un campione più eterogeneo e quindi composto da non soli studenti (Frederick, 2005). Stanovich e West (2000) hanno sostenuto che una maggiore comprensione delle informazioni può derivare da un intelletto superiore (così come da una grande capacità deliberativa o da punteggi più alti su WM a breve termine). In risposta a coloro che hanno sostenuto che le risposte comunemente classificate come errori sono in realtà delle risposte coerenti a diverse interpretazioni della domanda (per esempio, Hinton, 1995), Stanovich e West (2000) hanno risposto che, se gli intervistati più intelligenti sono più propensi a dare risposte canonicamente corrette, le altre risposte non possono essere considerate altrettanto buone. Di contro, alcuni respingono l'idea di una correlazione tra un certo grado di intelligenza (misurata come WM a breve termine) e la capacità di dare la "migliore" risposta (Sternberg, 2000). In risposta a questo, Bar Hillel (1991, pag 413) affermò che: "Molti scrittori hanno tentato di difendere le risposte apparentemente errate offrendo interpretazioni riguardo il ragionamento adottato da questi soggetti per razionalizzare le loro risposte. A volte, tuttavia, questo approccio caritatevole si è rivelato sbagliato in quanto, spesso i soggetti sono pronti a riconoscere il proprio errore una volta che gli è stato fatto notare, e perché l'interpretazione richiesta per giustificare la risposta errata è ancora più imbarazzante dell'errore stesso." È sorprendente notare come i soggetti non riconoscano l'errore prima della spiegazione, in quanto il "problema" è facilmente comprensibile, una volta chiarificato. Tuttavia, quando un problema è troppo difficile per tutti, la correlazione si può invertire in quanto gli intervistati più intelligenti sono più propensi a concordare un errore plausibile piuttosto che rispondere a caso (Kahneman, 2000b). Pertanto la difficoltà dei CRT (come è stato detto precedentemente) non è ad un livello tale da poter fare questo tipo di osservazioni.

Nel 1956, lo psicologo George Miller ha definito la memoria a breve termine come la facoltà della mente umana di poter contenere una quantità limitata di informazioni in uno stato accessibile. Ha suggerito inoltre che la memoria umana a breve termine abbia uno spazio di circa sette elementi (più o meno due). Lo spazio per la memoria a breve termine è definito come la più lunga lista di elementi che una persona può ricordare nel corretto ordine dopo un breve periodo di tempo. Essa agisce come un bloc-notes in cui si ha il richiamo provvisorio delle informazioni ma decade rapidamente e ha una capacità limitata. Inoltre, Miller et al. (1960) si riferisce alla memoria in quanto viene utilizzata per adottare un comportamento o prendere una decisione economica. Nel valutare la memoria a breve termine si è tenuto conto della capacità di mantenere le informazioni in uno stato attivo, anche in presenza di interferenza. Il termine "Working Memory" ha avuto una rilevanza diversa dopo che Baddeley e Hitch (1974) hanno dimostrato che un singolo modulo non poteva tenere conto di tutti i tipi di

memoria temporanea. Dopo la pubblicazione del lavoro di Baddeley et al. (1974) la Working Memory è stata considerata come la combinazione di più componenti che lavorano insieme e che possiede le stesse caratteristiche della memoria a breve termine, compreso il suo spazio. Ricerche più recenti hanno dimostrato che questo numero è approssimativamente accurato per gli studenti universitari, ma lo spazio della Working Memory a breve termine varia ampiamente con il tipo di popolazione testata, il materiale utilizzato (parole, numeri o immagini), il livello di carico cognitivo proposto e con il ricorso ad incentivi monetari reali (Cowan, 2001). La mia indagine sulla Working Memory a breve termine si focalizza sul ruolo dell'attenzione a breve termine, e sul fatto che la Working Memory a breve termine sia positivamente correlata con il numero di elementi ricordati. La capacità di controllare l'attenzione è definita come la limitata capacità di mantenere l'attenzione concentrata su una cosa e non lasciarsi distrarre da altri eventi, provenienti dall'esterno o dall'interno, generati da pensieri e sentimenti (Barret, Tugade e Engle, 2004). Molti studi indicano che la Working Memory a breve termine varia tra le persone, e quindi evidenzia le differenze individuali nelle capacità intellettuali, tra cui la capacità di controllare l'attenzione, e si modifica nell'arco della vita (Cowan, 2005; Engle, 2002). Maggiore Working Memory riflette maggiore capacità di controllare l'attenzione, specialmente in circostanze in cui vi siano interferenze o distrazione. La Working Memory dovrebbe essere correlata ad una serie di attività cognitive complesse come l'Effetto Interferenza e l'Effetto Soppressione. Effetto Interferenza: la capacità di mantenere attivamente le informazioni acquisite richiede risorse cognitive per non lasciarsi influenzare dalle distrazioni. La Working Memory a breve termine è legata a questa capacità di resistere alle interferenze, mantenendo attivamente le informazioni in linea con l'obiettivo (Barret, Tugade e Engle, 2004). Effetto Soppressione: la Working Memory a breve termine è legata alla capacità di controllare l'attenzione per mantenere alcune informazioni fuori dalla memoria (Barret, Tugade e Engle, 2004). Le prove di queste relazioni ci sono fornite sia da esperimenti economici che da esperimenti psicologici. Ci sono alcune indicazioni che vedono le differenze individuali nella WM a breve termine. C'è qualche indizio che vede, gli individui caratterizzati da differenti livelli WM a breve termine, differire per il loro uso del processo decisionale automatico e controllato durante le attività impegnative in laboratorio. Queste differenze nel comportamento portano a delle metafore molto diverse per descrivere il processo decisione di coloro che hanno una Working Memory a breve termine relativamente alta rispetto ai soggetti con una Working Memory a breve termine bassa. I soggetti con Working Memory a breve termine più elevata possono essere descritti con la metafora del "tattico motivato" (Fiske & Taylor, 1991). I tattici motivati hanno a disposizione molte risorse per elaborare le informazioni disponibili e possono

controllare efficientemente la loro attenzione sulla base dei loro obiettivi e motivazioni. Al contrario, i soggetti con una Working Memory a breve termine più bassa possono essere meglio descritti con la metafora degli “avari cognitivi” (Taylor, 1981). Questi soggetti con una Working Memory a breve termine inferiore, presentano delle risorse attentive fortemente limitate, e di conseguenza, adottano strategie che riducono la necessità di controllare l'attenzione. Anche se possono avere una serie di obiettivi o motivi, non hanno le risorse attentive sufficienti per mantenere le informazioni obiettivo rilevanti di fronte a situazioni complesse. Questa è solo una teoria ispirata dagli errori cognitivi nelle prove psicologiche in laboratorio. Senza incentivi monetari, Conway et al. (2001) trovarono che gli individui che avevano ottenuto un punteggio alto della Working Memory a breve termine, misurata attraverso l'ascolto di un suono per essere successivamente ricordato, hanno notato i loro nomi in un canale secondario di ascolto dicotico molto meno spesso di quanto non lo avessero fatto le persone con un basso punteggio di Working Memory a breve termine. Gli individui che avevano ottenuto un alto punteggio, a quanto pare, sono in grado di svolgere meglio il loro compito principale, essendo meno vulnerabili alle interferenze, ma questo va a discapito di essere un po' incurante riguardo gli aspetti irrilevanti dell'ambiente. In risposta a tale ricerca, Engle e i suoi colleghi attraverso la loro evidenza empirica senza incentivi reali hanno dimostrato che l'aumento della Working Memory a breve termine facilita l'elaborazione di una maggiore attenzione resistente alle interferenze provenienti da altre fonti e facilita il ricordo di più parole. In questo modo, il concetto di Working Memory a breve termine e i processi relativi al controllo dell'attenzione a breve termine si fanno sempre più interconnessi. Nel nostro esperimento si usa un modo migliore per misurare la Working Memory a breve termine. Le attività che misurano le differenze individuali nella capacità della Working Memory a breve termine sono alcuni degli strumenti più non standardizzati usati oggi in psicologia cognitiva (Conway et al., 2005). Anche se questi strumenti sono cruciali nella comprensione dei meccanismi e le relazioni sottostanti il processo decisionale, sono anche costantemente utilizzati nella letteratura economica più ampia, come predittivi di una vasta gamma di capacità cognitive. Ad esempio, gli individui con punteggi più alti su varie prove di abilità cognitive tendono ad essere meno soggetti a bias comportamentali e fallimenti nel ragionamento (ad esempio, Redick e Engle et al 2012;. Oechssler et al, 2009;. Stanovich e West, 2000;. Toplak et al, 2011). Perché questi strumenti li ritroviamo spesso in letteratura? Forse l'aspetto più importante per misurare la Working Memory a breve termine è che essa prevede un certo numero di caratteristiche fondamentali delle persone. Ad esempio, le persone con una Working Memory più alta tendono ad ottenere una performance maggiore nel multi-tasking, seguono meglio le indicazioni, hanno ottenuto

maggiori punteggi SAT e sono più intelligenti (Engle, 2002; Engle & Kane, 2004). Misurare correttamente le differenze individuali nella Working Memory a breve termine è abbastanza complicato, pertanto per rendere la misurazione precisa e affidabile, Engle e i suoi colleghi hanno creato e reso liberamente disponibile le versioni automatizzate di tre degli strumenti più utilizzati per misurare la Working Memory a breve termine (Operation, Symmetry, and Reading Span Task), utilizzando diverse parole come elementi da ricordare. Essi fanno parte della famiglia dei Complex Span Task (CST) e tengono conto delle considerazioni psicologiche e teoriche conosciute che influenzano i punteggi su questi Task (Redick and Engle et al. 2012). I CST impegnano la Working Memory a breve termine, in quanto i partecipanti sono tenuti a mantenere le informazioni attive e rapidamente recuperabili, nonostante la loro attenzione sia periodicamente spostata verso le attività di interferenza (Baddeley e Hitch, 1974). In questa tesi, usiamo l'Operation Span Task per dare un valore alla Working Memory a breve termine individuale, ponendo nei trattamenti un maggior carico cognitivo. Esso consiste di due attività ripetute e mostrate contemporaneamente attraverso varie schermate di un PC. I partecipanti hanno osservato sia una equazione aritmetica (come interferenza) da risolvere e sia una parola da ricordare e da richiamare poco tempo dopo. Una volta terminato un numero variabile di schermate, è stata presentata una griglia in cui i partecipanti hanno inserito le parole nel corretto ordine in cui sono state presentate (come proxy della Working Memory a breve termine). Pertanto, il punteggio della Working Memory a breve termine è misurato dalla percentuale di parole ricordate. Per spiegare l'ampio programma di utilità predittiva dei CST, Engle e colleghi hanno proposto che i risultati sul funzionamento di Operation, Symmetry, and Reading Span riflettano principalmente le differenze individuali nella capacità di controllare l'attenzione (Engle & Kane, 2004; Kane et al., 2007). Il controllo dell'attenzione prevede sia mantenere e sia recuperare le informazioni (Unsworth & Engle, 2007). La capacità di mantenere e recuperare le informazioni è particolarmente importante in situazioni in cui ci siano molteplici interferenze, come ad esempio i CST, dove l'attenzione deve passare per rappresentazioni mentali, in quanto le informazioni devono essere ricordate e poi rapidamente dimenticate di nuovo all'interno dell'esperimento. L'abilità di selezionare le informazioni (abilità nel controllo dell'attenzione) e di conservare queste informazioni di breve termine in uno stato accessibile (Working Memory a breve termine) sono aspetti critici delle nostre capacità cognitive (Barret, Tugade and Engle, 2004). Pertanto, assumiamo che la migliore Working Memory a breve termine è positivamente correlata con il maggiore numero di parole ricordate (attenzione). In altri termini, la nostra principale ipotesi congiunta è che il punteggio più alto di WM a breve

termine permette un controllo maggiore dell'attenzione attivando le funzioni del processo cognitivo razionale (Sistema 2).

Tuttavia, ci sono tre effetti principali che influenzano la nostra prestazione di Working Memory a breve termine: l'Effetto Prova, l'Effetto Lunghezza Parola e l'Effetto Ordine. Il primo migliora prestazioni mantenendo attive e in memoria circa 2-4 elementi aggiuntivi se i partecipanti hanno abbastanza tempo e se possono provare la "parola da richiamare" in mente in silenzio (Campbell & Bagshaw, 2008). La lunghezza delle parole ha dimostrato essere una variabile molto rilevante nelle attività della Working Memory a breve termine (Baddeley, 1986). Baddeley, Thomson e Buchanan (1975) hanno dimostrato, per esempio, che i soggetti ricordano più le parole brevi che quelle lunghe. L'Effetto Ordine si verifica quando devono essere tenuti nella Working Memory a breve termine diversi elementi (come cifre, parole o immagini) nello stesso tempo, questi si trovano a competere effettivamente con gli altri elementi per il richiamo. I nuovi elementi, di conseguenza, a poco a poco spingono le informazioni più vecchie, a meno che gli elementi più vecchi non siano attivamente protetti contro le interferenze dell'Effetto Prova o dal controllo dell'attenzione (Turner e Engle, 1989). Di conseguenza se l'ordine delle attività è sempre lo stesso (tre parole da richiamare, poi quattro e in ultimo cinque parole da richiamare), i soggetti possono apprendere dalle prime attività e, successivamente, ricordare più parole (Redick e Engle et al., 2012). Per la quantità di tempo e provando le parole, potremmo affermare che essi non abbiano abbastanza tempo, ma non possiamo affermare con sicurezza che sia così. Questo effetto è usato raramente nelle attività ad alto livello cognitivo come CST (Redick e Engle et al., 2012). Al fine di verificare l'ipotesi dell'Effetto Lunghezza Parole, prendiamo in considerazione una variabile dummy (come Baddeley, Thomson, e Buchanan (1975); La Point e Engle, 1990). Mentre per quanto riguarda l'Effetto Ordine, abbiamo il sospetto che influenzi i nostri risultati. Ci sono anche altri tipi di attività, chiamati semplicemente span tasks, che sono stati ampiamente utilizzati in una serie di test di intelligenza standardizzati. In attività semplici (quali Digit Task, Word Task and Corsi Blocks), i partecipanti riportano in maniera sequenziale una serie di elementi presentati. Più specificamente, una singola semplice attività misura non solo la capacità cognitiva in questione, ma anche altri fattori estranei a quella capacità, come la velocità a risolvere dei problemi matematici, o la capacità di lettura, ecc (Loehlin 2004; Wittman, 1988). In altre parole, un unico indicatore di Working Memory non può essere considerato come misura di una Working Memory a breve termine. Pertanto, la famiglia dei CTS differisce dai semplici span task per l'aggiunta di una attività di interferenza (utilizza almeno due paradigmi). Nei CST vengono codificati degli elementi di memoria (ad esempio, parole) alternati con attività secondarie che richiedono attenzione (ad esempio, la lettura di frasi

o / e la risoluzione di un problema matematico). Dunque, per trarre delle specifiche conclusioni circa la Working Memory a breve termine i ricercatori dovrebbero utilizzare più indicatori per creare un punteggio composito che consista della varianza condivisa tra Working Memory e due o più paradigmi (Conway et al, 2005;. Engle et al, 1999;. Loehlin 2004; Wittman, 1988). Pertanto, l'affidabilità del CST è stata ripetutamente verificata (Redick e Engle et al., 2012). Si potrebbe criticare il fatto che in questo studio abbiamo utilizzato un solo tipo di CST: l' Operation Span Task. Tuttavia, Kane et al. (2004) hanno dimostrato che i componenti di Reading, Operation, and Symmetry Span tengono conto di una varianza simile nelle prove di abilità verbali e spaziali, che porta a ritenere che le differenze individuali nella WM a breve termine siano di dominio generale. In quanto tale, l'utilizzo di più paradigmi per misurare la Working Memory a breve termine dovrebbe tradursi in una misurazione migliore rispetto alle multiple attività provenienti tutte dallo stesso paradigma. Per tale ragione, in questo studio, usiamo una Complex Span Task dove ci sono due paradigmi (menzionati sopra) che creano un punteggio composito della Working Memory a breve termine. Un fattore importante, che ha ricevuto scarsa attenzione nella letteratura, riguarda la struttura degli incentivi in base al quale i soggetti vengono ricompensati per ogni risposta corretta nell'esperimento. E' ancora un campo aperto quello relativo all'interazione tra i bias cognitivi ed i sistemi di incentivazione. Questo problema è stato affrontato prima da Chamberlin (1948), che ha suggerito un metodo per risolverlo, essenzialmente fornendo ai soggetti incentivi reali monetari. Ad esempio, è stato dimostrato che i soggetti avevano un rendimento migliore sui CRT se questi prevedevano incentivi monetari. Ponti, Cueva et al. (2015) dando un'occhiata ai CRT privi di incentivi constatarono un punteggio medio di 1,08 per i maschi e 0,55 per le femmine. Così come Bosch e Domenech et al. (2014), in assenza di incentivi, trovarono punteggi medi di 0,95 e 0,58 per i maschi e le femmine. Un altro è il lavoro di Altmann, Falk e Grunewald (2013) in cui i punteggi medi erano di 0,8 senza differenze di genere e di pagamenti specifici per la risposta corretta ai CRT. D'altra parte Frederick (2005), usando incentivi monetari, riscontrò un punteggio medio di 1,47 per i maschi e 1,03 per le femmine. Oechssler et al. (2009) hanno osservato CRT prevedendo incentivi monetari trovò un punteggio medio di 2,2 per i maschi e 1,7 per le femmine. Hoppe e Kusterer (2011) fecero la stessa cosa ed evidenziarono punteggi di 2.12 e 1.61 rispettivamente per i maschi e le femmine. Come è chiaro da questi risultati, vi è un forte Effetto Genere (i maschi generalmente raggiungono un punteggio significativamente più alto rispetto alle femmine), ma è evidente che le differenze di genere sono ridotte quando il CRT prevede un incentivo monetario (Ponti e Cueva et al. (2015). Nel 1976, Smith sottolineò l'importanza di fornire ai soggetti incentivi monetari sufficienti, al fine di superare gli effetti

distorsivi dei "costi cognitivi decisionali", cioè l'impegno nella riflessione. Gli economisti in genere danno per scontato che il processo decisionale sia motivato principalmente da incentivi materiali, e che le decisioni economiche siano regolate principalmente da interessi e razionalità. In questo contesto, la razionalità significa che coloro che devono prendere una decisione utilizzano le informazioni disponibili in modo logico e sistematico, in modo da fare scelte ottimali dato le alternative a portata di mano e l'obiettivo da raggiungere (Kahneman, 2002). Secondo questo punto di vista, in questo esperimento sono stati utilizzati degli incentivi monetari reali sia per il Complex Span Task che per i CRT. Questa tesi è una delle prime che tenta di spiegare questi legami attraverso un esperimento economico, oltre ad essere il primo tentativo di mettere in relazione la WM a breve termine per le prestazioni cognitive coinvolgendo degli incentivi reali.

Design dell'esperimento

Ipotesi

In questa tesi esaminiamo le seguenti quattro ipotesi:

1. Le Performance nei CRT migliorano con la Working Memory a breve termine misurata attraverso le parole da ricordare.
2. L'Effetto di Genere deteriora le performance delle femmine nei CRT.
3. Il numero di parole ricordate (come proxy della WM a breve termine) si riduce con l'aumentare del numero delle parole da mantenere in memoria.
4. Il numero di parole ricordate (come proxy della WM a breve termine) si riduce con l'aumentare della lunghezza delle parole da mantenere in memoria.

La nostra **Ipotesi 1** è stata analizzata nello studio di Barret, Tugade e Engle (2004). Rappresenta un'ipotesi congiunta in quanto la letteratura teorica ipotizza che la Working Memory a breve termine è correlata con l'attenzione controllata che guida i partecipanti verso l'utilizzo del processo cognitivo deliberativo il quale migliora le performance nei CRT. Rigettare questa ipotesi potrebbe significare il rigetto di uno qualsiasi (o più di uno) degli elementi della teoria, quali:

- (i) La WM a breve termine può non aumentare l'attenzione controllata.
- (ii) L'attenzione non guida i partecipanti verso lo stato cognitivo deliberativo.
- (iii) Lo score nei CRT non migliora quando si usa lo stato cognitivo deliberativo.

La nostra **Ipotesi 2** è stata analizzata dalla letteratura e da Frederick (2005), e si fonda sulla forte componente di “Effetto Genere” (Maschio, Femmina) sulle prestazioni del CRT. I maschi hanno un punteggio superiore rispetto alle femmine. Rigettare questa ipotesi significa che non c'è alcuna differenza nelle performance nei CRT tra maschi e femmine.

La nostra **Ipotesi 3** è stata analizzata nello studio di Turner e Engle (1989), e si fonda sul deterioramento della Working Memory a breve termine con l'aumentare del numero di parole da tenere in mente. Rigettare questa ipotesi significa che l'aumento del numero delle parole da ricordare non influenza la percentuale di parole ricordate.

La nostra **Ipotesi 4** è stata analizzata nello studio di Baddeley, Thomson e Buchanan (1975), e si fonda sul deterioramento della Working Memory a breve termine con l'aumentare della lunghezza delle parole da ricordare. Rigettare questa ipotesi significa che l'aumento della lunghezza delle parole da tenere in mente non influenza la percentuale di parole ricordate.

Partecipanti

I partecipanti al nostro esperimento sono 134 studenti della University of Central Florida negli USA, di età compresa tra 18 e 36 con una media di 22.

Procedura

L'esperimento è stato condotto nell'arco di diversi mesi tra il 2007 e il 2008 per un totale di 12 sessioni. E' composto da due Task da completare ordinatamente uno alla volta. Tutti gli input di ciascun partecipante sono stati registrati sul computer. Prima di iniziare, i partecipanti hanno ricevuto le adeguate istruzioni di ogni attività. I partecipanti non hanno potuto utilizzare altri PC, tranne quello fornito dagli sperimentatori. Una volta che tutti sono stati registrati nel programma, l'esperimento è stato avviato attraverso un monitor centrale nella stanza. Al termine dell'esperimento ogni partecipante è stato pagato in privato.

Task 1: Operation Span Task

In questa sezione, usiamo un totale di 36 attività divise per 9 trattamenti per misurare la Working Memory a breve termine. Ogni attività è stata proiettata sullo schermo del PC personale, dividendolo orizzontalmente in due parti parallele. Nella parte alta è stata presentata un'equazione matematica chiedendo di indicare se tale equazione fosse corretta oppure no (come attività di interferenza). Mostriamo un esempio di un'equazione matematica usata:

$5 * 5 - 10 = 15 ?$ Yes No

Nella parte bassa dello schermo viene esposta una parola per il successivo richiamo (come proxy della Working Memory a breve termine). La lunghezza di tale parola varia a seconda della sessione.

Abbiamo diviso i 9 trattamenti in tre gruppi, ognuno contenente tre trattamenti. Nel primo gruppo sono state mostrate ad ogni trattamento tre schermate come quella descritta precedentemente, e alla fine è stata presentata una griglia per inserire le parole esposte poco tempo prima (per un totale di 9). Successivamente, nell'altro gruppo di tre trattamenti le schermate mostrate ad ogni trattamento sono diventate quattro, ed anche in questo caso alla fine dell'ultimo trattamento è stata esposta la relativa griglia da compilare con le parole (per un totale di 12). Nell'ultimo gruppo ne sono state mostrate cinque ad ogni trattamento, con l'ultimo trattamento seguito dalla griglia da compilare con le parole (per un totale di 15). Dopo aver completato il totale di 36 equazioni matematiche e parole da ricordare questo Task volge al termine. I partecipanti sono stati pagati 25 centesimi per ogni equazione matematica corretta e 25 centesimi per ogni parola ricordata correttamente (o compitata correttamente). Hanno potuto guadagnare un massimo di 18 \$ nel Task 1.

Task 2: Cognitive Reflection Test e Domande Demografiche

Questa è la sezione finale del nostro esperimento. Dopo aver completato il Task 1, il computer mostra un'ultima schermata in cui i tre CRT sono posti in cima allo schermo seguiti in basso dalle domande demografiche. Il Task 2 è formato da 22 domande, le cui prime tre sono i problemi esposti nei CRT e i partecipanti sono valutati e pagati 1\$ per ogni risposta corretta (come proxy che stabilisce l'efficiente utilizzo del processo cognitivo deliberativo). Mentre i partecipanti non hanno ricevuto nessun pagamento per le altre 19 domande demografiche, ma presumiamo che abbiano risposto il più sinceramente possibile.

Statistiche Descrittive

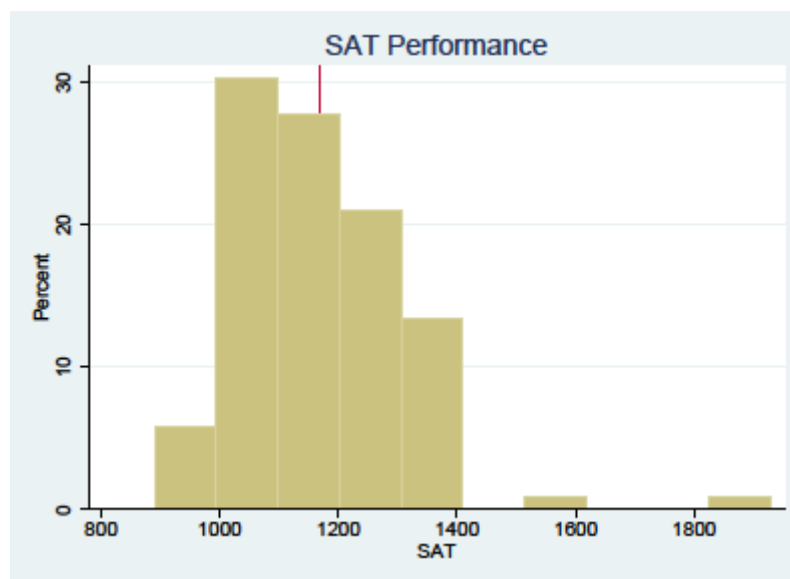
I dati sono stati raccolti tra il 2007 e il 2008 e i partecipanti erano 134 studenti della University of Central Florida, di età compresa tra 18 e 36 con una media di 22. Più del 91% dei nostri partecipanti aveva una media di 25 anni o inferiore, sottolineando un range di età basso. Come prevedibile, quasi tutti sono cittadini USA, single o non sposati. Più dei due terzi dei partecipanti sono di etnia bianca, e più di un quarto dell'intero campione è di religione cattolica. Sono stati presi in considerazione anche il reddito e il livello di istruzione dei genitori, i quali sono risultati medio-alti per la maggioranza dei partecipanti. Per ultimo, è stato riscontrato che quasi nessuno fuma o fuma meno di un pacchetto di sigarette al giorno. Da questa prima analisi possiamo dire che abbiamo un'eterogeneità della popolazione molto limitata, in quanto tutti i partecipanti sono giovani studenti che sono stati reclutati nella stessa università. Nel nostro esperimento, i partecipanti non coprono un ampio raggio di abilità cognitive in quanto appartengono tutti alla stessa classe di abilità cognitive chiamata "college status" da Redick e Engle (2012). Pertanto, se un ricercatore è interessato a trarre conclusioni generali o specifiche riguardo la Working Memory a breve termine può arrivare a risultati differenti se i partecipanti all'esperimento sono tutti studenti rispetto a una popolazione campionaria in cui sono compresi anche i non-studenti. Sulla stessa linea, Redick e Engle (2012) trovarono che gli studenti ottenevano un punteggio leggermente superiore rispetto ai non-studenti nello svolgimento dell'Operation Span Task senza incentivi reali. Hanno inoltre esaminato la performance sulla base dell'università di frequenza dei partecipanti (University of Georgia (UGA), University of North Carolina Greensboro (UNCG) e Georgia Institute of Technology (GT)). I punteggi dell'Operation Span Task computerizzato degli studenti GT sono stati leggermente superiori per quelli degli studenti UGA, ed entrambi hanno ottenuto un punteggio superiore rispetto agli studenti UNCG. I punteggi degli studenti UNCG sono stati più simili al gruppo di partecipanti non-studenti, rappresentato da un mix di studenti universitari della zona di Atlanta e persone non studenti.

La stessa analisi è stata fatta sui CRT. Il punteggio dei tre CRT è stato calcolato dando ai partecipanti un punto per ogni risposta corretta. Garza, Kujal e Lenkei (2015) mostrarono, con incentivi monetari, che gli studenti ottengono un punteggio sostanzialmente superiore rispetto ai non-studenti nel CRT 1, mentre solo leggermente maggiore nelle performance del CRT 2 e del CRT 3. Inoltre, trovarono che gli studenti universitari hanno meno probabilità di dare la risposta sbagliata in tutte e tre le domande, e di conseguenza hanno maggiore probabilità di

dare almeno due risposte corrette su tre rispetto ai non-studenti. L'Effetto Genere sembra influenzare allo stesso modo studenti e non-studenti, dando un vantaggio nelle performance ai maschi. In conclusione possiamo affermare che i nostri risultati, sia sulla Working Memory a breve termine che sui CRT, sono influenzati dal "college status".

Dopo aver definito il problema del nostro esperimento, descriviamo le qualità accademiche dei nostri partecipanti. Più dei due terzi frequentavano l'ultimo o il penultimo anno dell'università e la metà dei partecipanti pensava di lasciare gli studi dopo aver conseguito la Laurea Magistrale. Più del 75% dei partecipanti ha dichiarato di avere un GPA uguale o superiore a "B" e più della metà dei partecipanti segue come corso principale di studio quello inerente all'economia e alla finanza (materie matematiche). Infine, la media dei punteggi del SAT dei partecipanti è stata di 1168. Tutti questi dati sono stati richiesti in un questionario demografico durante l'esperimento.

Grafico 1. Punteggio SAT dei nostri partecipanti richiesto nel questionario demografico



Il Grafico 1 mostra come i valori del SAT dei nostri partecipanti sono concentrati intorno alla media (1168). Redick e Engle (2012), in accordo con Princeton Review, hanno mostrato che i punteggi del SAT degli studenti sono sempre concentrati intorno alla media. In quanto tale media cambia in base alla università stessa (ad esempio i punteggi medi del SAT in: GT (1334),

UGA (1234) e UNCG (1045)). In ogni università i valori medi del SAT degli studenti cambiano ma restano concentrati intorno alla loro media, confermando i nostri risultati.

Tab. 1 Relazione tra CRT Score e GPA, SAT, Percentuale di parole ricordate

<i>Correlazione</i>	CRT score	% Parole Ricordate	GPA
CRT score			
% Parole Ricordate	23.13%		
GPA	17.19%	13.71%	
SAT	41.48%	21.39%	15.41%

La Tabella 1 mostra una correlazione positiva tra la percentuale delle parole ricordate correttamente e il punteggio SAT nel nostro esperimento, confermando i risultati di Redick e Engle (2012). Ciò significa che il punteggio SAT è una funzione “dell’attenzione” che stiamo assumendo correlata con la percentuale di parole richiamate. Inoltre come supporto alla nostra ipotesi principale, vediamo che il punteggio dei CRT e la percentuale di parole ricordate correttamente hanno una relazione positiva. In linea con Frederick (2005) e Ponti & Cueva (2015), vediamo anche alcune evidenze sui punteggi dei CRT che sono altamente e positivamente correlati con il Scholastic Achievement Test (SAT) e con la media dei voti GPA. Pertanto, i punteggi elevati su SAT e GPA danno benefici in termini di maggiore punteggio sui CRT, così come una maggiore percentuale di parole ricordate correttamente.

Differenze di genere

Una quantità equilibrata di partecipanti maschili e femmili ha aderito al nostro esperimento.

Per quanto riguarda l'Operation Span Task, precedenti studi di Redick e Engle (2012) hanno indicato che i maschi ricordano un numero maggiore di parole rispetto alle femmine, senza incentivi monetari. Il vantaggio maschile è stato solamente di 1-2 parole, e come indicano le dimensioni dell'effetto, l'entità dell'Effetto Genere è molto basso. Pertanto è difficile affermare che la Working Memory a breve termine sia superiore per i maschi piuttosto che per le femmine. In linea con questi studi, abbiamo trovato in media una differenza nella prestazione di massimo 2 parole ricordate in più per i maschi. Tuttavia, abbiamo testato questa ipotesi utilizzando il Mann-Whitney Test³ trovando la differenza tra maschi e femmine non significativa e con un p-value di 0,1058⁴. Inoltre, l'Effetto Genere non è statisticamente significativo nelle prestazioni aritmetiche a causa di un p-value molto elevato (0,7034)⁵. Quindi, nell'Operation Span Task accettiamo ipotesi nulla in cui non ci sono differenze nelle prestazioni della WM a breve termine tra maschi e femmine.

Per continuare l'analisi su questo effetto, andiamo a indagare sulle differenze individuali nelle prestazioni sui CRT. Generiamo una variabile che assegna ai partecipanti un punto per ogni risposta corretta sui CRT. E' la nostra proxy che stabilisce l'efficiente utilizzo del processo cognitivo deliberativo. Nella letteratura, i CRT hanno guadagnato attenzione anche grazie al fatto che, contrariamente alle altre proxy delle abilità cognitive come il SAT o i Complex Span Task, le femmine ottengono uno score significativamente più basso rispetto ai maschi. L'Effetto Genere sui CRT è stato mostrato in numerosi studi. Con un'elevata numerosità campionaria (3428), Frederick (2005) ha trovato che i maschi raggiungono mediamente punteggi più elevati rispetto alle femmine (vedi anche Oechssler et al. 2009, Hoppe & Kusterer 2011; Ponti, Cueva et al 2015, Bosch & Domenech et al. 2014; Garza, Kujal & Lenkei, 2015). In questi lavori hanno dimostrato che:

(i) i maschi hanno più probabilità di rispondere correttamente in tutte e tre le domande prese singolarmente.

(ii) le femmine hanno più probabilità di rispondere a nessuna di queste correttamente.

³ Il Mann-Whitney Test considera una distribuzione più generale della popolazione rispetto a una distribuzione normale (T-Test).

⁴ Il t-test ha un p-value di 0.0471.

⁵ Il t-test ha un p-value di 0.3450.

(iii) i maschi hanno più probabilità di rispondere a tutte e tre le domande correttamente.

È importante sottolineare che le differenze di genere persistono anche quando le si controlla cambiando le caratteristiche del test (ad esempio utilizzando incentivi monetari, computer, variando la popolazione campionaria studentesca e il posizionamento dei CRT nell'esperimento, vedi Garza, Kujal e Lenkei, 2015).

Tab. 2 **Effetto Genere sul CRT Score**

CRT1	Gender		Tot.
	M	F	
0	46	43	89
	59.7%	75.4%	
1	31	14	45
	40.3%	24.6%	
Tot.	77	57	134

CRT2	Gender		Tot.
	M	F	
0	53	42	95
	68.8%	73.7%	
1	24	15	39
	31.5%	26.3%	
Tot.	77	57	134

CRT3	Gender		Tot.
	M	F	
0	34	33	67
	44.2%	57.9%	
1	43	24	67
	55.8%	42.1%	
Tot.	77	57	134

Global CRT Score	Gender		Tot.
	M	F	
0	24	29	53
	31.1%	50.9%	
1	22	9	31
	28.6%	15.8%	
2	17	13	30
	22.2%	22.8%	
3	14	6	20
	18.1%	10.5%	
Tot.	77	57	134

Le celle gialle rappresentano le percentuali di Genere.

Complessivamente, i nostri risultati confermano una forte componente di Effetto Genere nelle performance dei CRT. Il CRT score riceve un punto per ogni risposta corretta e zero punti altrimenti ed è la nostra proxy che manifesta la prevalenza del processo cognitivo deliberativo.

La Tabella 2 verifica le tre sub-ipotesi menzionate sopra:

(i) in tutti e tre i CRT è stata più alta la percentuale di risposta corretta per i maschi piuttosto che per le femmine⁶.

(ii) il 50.9% delle femmine non ha risposto correttamente a nessuna delle tre domande, contro il 31.3% dei maschi⁷.

(iii) il 18.1% dei maschi ha dato la risposta corretta in tutti e tre i CRT, mentre solo il 10.5% delle femmine ha fatto lo stesso⁸.

Per esempio, anche Garza, Kujal and Lenkei, (2015) nel loro lavoro hanno trovato che in tutte e tre le domande prese singolarmente è più alta la probabilità di risposta corretta per i maschi piuttosto che per le femmine. Hanno mostrato che la grande maggioranza dei partecipanti, i quali non hanno risposto correttamente a tutti e tre i quesiti, è rappresentata dalle femmine. Inoltre, usando il criterio delle tre risposte corrette, hanno indicato che la probabilità per i maschi era il doppio rispetto a quella delle femmine. Ulteriori prove ce le ha fornite Frederick nel suo lavoro del 2005, in cui ha diviso i partecipanti in due gruppi, uno costituito dai partecipanti che hanno risposto solo correttamente nei CRT (High), l'altro costituito da chi ha dato solo risposte sbagliate (Low). Ha dimostrato che la performance dei maschi è più alta rispetto a quella delle femmine, in quanto il gruppo "High" è costituito per i due terzi da maschi, mentre il gruppo "Low" è costituito per i due terzi da femmine. L'autore ha aggiunto che le disparità di prestazione nei CRT tra questi due gruppi possono rivelare altre differenze di genere oltre la riflessione cognitiva.

In Appendice A è mostrata visivamente la differenza di genere nella prestazione dei tre CRT. In una scala da 0 a 3 punti, in media, nel nostro esperimento è stato ottenuto un punteggio di 1.13 e lo score per i maschi è stato di 1.27 mentre per le femmine di 0.93. Sulla stessa linea di Frederick (2005), il quale ha utilizzato incentivi monetari per le prestazioni e ha verificato una media globale di 1.24 suddivisa in 1.47 per i maschi e 1.03 per le femmine. Anche Hoppe e

⁶ Il Mann-Whitney test ha un p-value di 0.0581 per il CRT 1, 0.0924 per il CRT 2 e 0.0772 per il CRT 3, mentre il t-test ha un p-value di 0.0289 per il CRT 1, 0.2722 per il CRT 2 e 0.0588 per il CRT 3.

⁷ Il Mann-Whitney test ha un p-value di 0.0216, mentre il t-test ha un p-value di 0.0105

⁸ Il Mann-Whitney test ha un p-value di 0.0475, mentre il t-test ha un p-value di 0.1110

Kusterer (2011) hanno analizzato i CRT usando gli incentivi reali e hanno trovato una media di 1.32 per i maschi e 0.98 per le femmine. Analizzando i CRT senza incentivi monetari per le prestazioni, Ponti e Cueva (2015) hanno osservato delle medie leggermente inferiori e pari a 1.12 per i maschi e 0.58 per le femmine. Quindi, è stato verificato un leggero miglioramento nella performance di entrambi i generi attraverso il ricorso agli incentivi monetari reali. Abbiamo testato tale differenza nelle prestazioni maschio/femmina usando il Mann-Whitney Test e abbiamo trovato un Effetto Genere statisticamente significativo con un p-value di 0.0574⁹. Questo sarebbe coerente con l'interpretazione che le femmine sono più impulsive rispetto ai maschi e vogliamo anche suggerire che in media i maschi sono più propensi ad usare il processo cognitivo deliberativo rispetto alle femmine. Pertanto, i maschi ottengono punteggi più alti nei CRT rispetto a quelli delle femmine.

Per confermare questa ipotesi, abbiamo esteso l'analisi usando un modo alternativo di lavorare i dati attraverso una nuova variabile dipendente chiamata "CTR Score Impulsivo". Lo Score Impulsivo è costruito assegnando un punto, per ogni CRT, se la risposta è definita come "risposta impulsiva" e zero punti altrimenti. Le risposte impulsive sono 10 per il CRT 1, 100 per il CRT 2 e 24 per il CRT 3. Tra tutte le possibili risposte sbagliate che la gente possa dare, le postulare risposte impulsive (10, 100 e 24) dominano. Infatti il CRT Score Impulsivo è costituito solamente dalle risposte impulsive e non da tutte le risposte sbagliate. Quindi, i CRT inducono una trappola cognitiva associata all'utilizzo del processo cognitivo euristico piuttosto che a quello deliberativo. In altre parole, lo Score Impulsivo sui CRT è inteso come la misura dell'incapacità di sopprimere le risposte impulsive erranee (Ponti, Cueva et al. 2015). Dovremmo ottenere risultati opposti alle precedenti statistiche descrittive, in quanto le risposte impulsive manifestano la prevalenza del processo cognitivo euristico nei partecipanti, il quale li induce in errore.

Tab. 3 Relazione tra CRT Score Impulsivo e GPA, SAT, Percentuale di parole ricordate

<i>Correlazione</i>	SAT	GPA	% Parole Ricordate
CRT score Impulsivo	-37.83%	-19.93%	-22.03%

Prima di tutto, la Tabella 3 mostra che la proxy per il processo cognitivo euristico si muove in direzione inversa rispetto a quando il CRT Score rappresentava le domande corrette. In quanto,

⁹ Il t-test ha un p-value di 0.0371.

un maggiore punteggio nel SAT o nel GPA o una maggiore percentuale di parole ricordate riduce l'effetto di dare la risposta impulsiva.

Abbiamo riscontrato una forte componente di Effetto Genere nella nostra proxy per il processo cognitivo deliberativo, quindi andiamo a verificare tale componente anche nella nostra nuova variabile che manifesta, invece, la prevalenza del Sistema 1 nei partecipanti. Le tre sub-ipotesi speculari sono:

- (i) le femmine hanno più probabilità di cadere nella trappola in tutti i CRT presi singolarmente.
- (ii) i maschi hanno più probabilità di rispondere a nessuno dei CRT impulsivamente.
- (iii) le femmine hanno più probabilità di rispondere impulsivamente a tutte e tre le domande.

Tab. 4 Effetto Genere sul CRT Score Impulsivo

Impulsive CRT1	Gender		Tot.
	M	F	
0	32	15	47
	41.6%	26.4%	
1	45	42	87
	58.4%	73.6%	
Tot.	77	57	134

Impulsive CRT2	Gender		Tot.
	M	F	
0	36	19	55
	46.7%	33.3%	
1	41	38	79
	53.3%	66.7%	
Tot.	77	57	134

Impulsive CRT3	Gender		Tot.
	M	F	
0	47	30	55
	61%	52.6%	
1	30	27	79
	39%	47.4%	
Tot.	77	57	134

Impulsive Global CRT Score	Gender		Tot.
	M	F	
0	17	9	26
	22.1%	15.8%	
1	23	13	36
	29.9%	22.8%	
2	18	11	29
	23.4%	19.3%	
3	19	24	43
	24.6%	42.1%	
Tot.	77	57	134

La Tabella 4 verifica le nostre sub ipotesi menzionate sopra:

(i) in tutti e tre i CRT è stata più alta la probabilità di risposta impulsiva per le femmine piuttosto che per i maschi¹⁰.

(ii) il 22.1% dei maschi non è caduto in nessuna delle risposte impulsive, mentre solo il 15.8% delle femmine fa lo stesso¹¹.

(iii) le femmine hanno avuto quasi il doppio della probabilità rispetto ai maschi di dare le tre risposte impulsive, in particolare il 42.1% delle femmine contro il 24.6% dei maschi¹².

Se lo Score Impulsivo dei CRT stabilisce il totale fallimento delle funzioni del Sistema 2 nel correggere la risposta impulsiva, sembrerebbe che i maschi manifestino spesso la prevalenza del processo cognitivo deliberativo, almeno in parte, in quanto solo il 24.6% di loro è stato totalmente impulsivo. In altre parole, le femmine hanno maggiori probabilità di essere divise quasi perfettamente in due gruppi: uno (il 42,1% delle femmine), costituito da chi ha risposto sempre impulsivamente, con conseguente fallimento totale delle funzioni di monitoraggio del Sistema 2 (confermato dal fatto che la metà delle femmine ha dato zero risposte corrette, vedere Tab. 2); e uno (il 57,9%), costituito da chi è riuscito a non dare almeno una risposta impulsiva, con conseguente parziale efficienza del processo cognitivo deliberativo.

Inoltre, se si confronta la Tabella 2 con la Tabella 4 possiamo fare un'ulteriore verifica per la validità della nostra ipotesi sull'Effetto Genere. Nel CRT 1 troviamo che le differenze di genere sono ridotte rispetto agli altri CRT, in quanto il 98% di entrambi maschi e femmine che hanno commesso l'errore in questa domanda, hanno dato la risposta impulsiva¹³. Inoltre, il 97% dei maschi che non è caduto nella trappola della risposta impulsiva, ha dato la risposta corretta contro il 90% delle femmine¹⁴. Nel CRT 2 troviamo che il 77,5% dei maschi che ha sbagliato questa domanda ha anche dato la risposta impulsiva contro il 90,5% delle femmine¹⁵. Inoltre qui notiamo una generale minore efficienza delle funzioni di controllo del Sistema 2, in quanto il 68% dei maschi che non ha risposto impulsivamente ha dato la risposta corretta, contro il 78% delle femmine¹⁶. Nel CRT 3 notiamo una generale minore impulsività, in quanto solo il

¹⁰ Il Mann-Whitney test ha un p-value di 0.0686 per il CRT1, 0.0987 per il CRT2 e 0.0823 per il CRT3, mentre il t-test ha un p-value di 0.0342 per il CRT1, 0.0601 per il CRT2 e 0.1671 per il CRT3.

¹¹ Il Mann-Whitney test ha un p-value di 0.0946, mentre il t-test ha un p-value di 0.1433

¹² Il Mann-Whitney test ha un p-value di 0.0333, mentre il t-test ha un p-value di 0.0164

¹³ Il Mann-Whitney Test ha un p-value di 0.7650, mentre il t-test ha un p-value di 0.5691

¹⁴ Il Mann-Whitney test ha un p-value di 0.0965, mentre il t-test ha un p-value di 0.1497

¹⁵ Il Mann-Whitney test ha un p-value di 0.0464, mentre il t-test ha un p-value di 0.0186

¹⁶ Il Mann-Whitney test ha un p-value di 0.0784, mentre il t-test ha un p-value di 0.0998

65% dei maschi che ha sbagliato in questa domanda ha dato la risposta impulsiva, contro il 64% delle femmine¹⁷. Inoltre, il 92% dei maschi che non ha risposto impulsivamente ha dato la risposta corretta, contro l'80% delle femmine¹⁸.

Tuttavia, l'analisi di queste percentuali in generale non confermano completamente la nostra ipotesi. Per esempio, la relazione percentuale tra chi ha dato la risposta impulsiva su chi non ha risposto correttamente ci indica sia il totale fallimento delle funzioni di controllo del processo cognitivo deliberativo, sia la prevalenza del processo cognitivo euristico. Nel nostro esperimento, in due su tre CRT questa prima relazione si è rivelata simile tra maschi e femmine, in quanto la differenza è risultata statisticamente non significativa al livello del 10%. Solo nel CRT 2 abbiamo trovato questa differenza percentuale statisticamente significativa, confermando la maggiore impulsività delle femmine. Invece, l'altra relazione percentuale tra chi ha dato la risposta esatta su chi non ha risposto impulsivamente ci indica l'efficiente funzione di monitoraggio del Sistema 2. In due su tre CRT questa seconda relazione ha confermato la nostra ipotesi, stabilendo la maggiore prevalenza ed efficienza del processo cognitivo deliberativo per i maschi. Nel CRT 2, invece, abbiamo ottenuto il risultato opposto, in quanto l'efficienza del Sistema 2 si è rivelata maggiore per le femmine piuttosto che per i maschi.

Facendo un'analisi più generale dividiamo maschi e femmine in quattro gruppi "estremi" nel senso che: uno ci indica la percentuale di maschi e di femmine che ha dato le tre risposte impulsive, uno ci indica la percentuale di chi ha dato zero risposte impulsive, uno ci indica la percentuale di chi ha dato le tre risposte corrette, e l'ultimo intuitivamente ci indica la percentuale di chi ha dato zero risposte corrette. Il 70.2% dei maschi che non ha dato le tre risposte corrette ha invece risposto impulsivamente a tutte e tre le domande, contro l'83.8% delle femmine¹⁹. Questo significa che le femmine le quali hanno commesso tre errori nei CRT hanno più probabilità di cadere nella trappola delle tre risposte impulsive manifestando una prevalenza del processo cognitivo euristico rispetto ai maschi. Mentre l'82% dei maschi che non ha dato le tre risposte impulsive ha invece risposto correttamente a tutte e tre le domande, contro il 66.4% delle femmine²⁰. In effetti, i maschi che non danno le tre risposte impulsive sono più propensi a manifestare un'efficienza del processo cognitivo deliberativo dando con maggiore probabilità le tre risposte corrette rispetto alle femmine. Pertanto, in generale, sembra

¹⁷ Il Mann-Whitney test ha un p-value di 0.3569, mentre il t-test ha un p-value di 0.2355

¹⁸ Il Mann-Whitney test ha un p-value di 0.0563, mentre il t-test ha un p-value di 0.0822

¹⁹ Il Mann-Whitney test ha un p-value di 0.0887, mentre il t-test ha un p-value di 0.1257

²⁰ Il Mann-Whitney test ha un p-value di 0.0385, mentre il t-test ha un p-value di 0.0912

corretto affermare che le femmine sono più impulsive rispetto ai maschi. Ma sembra ancora più corretto aggiungere che i maschi manifestano più spesso l'efficienza della funzione di controllo del Sistema 2, confermando la nostra ipotesi. Inoltre, anche se ci si concentra solo sui partecipanti che hanno dato le risposte sbagliate, maschi e femmine si differenziano. Gli errori delle femmine tendono a concentrarsi nella varietà impulsiva, mentre i maschi fanno una più ampia varietà di errori. Ad esempio, le femmine che hanno sbagliato il CRT 2 quasi sempre hanno dato la risposta impulsiva erronea di "100", mentre una frazione modesta dei maschi ha dato risposte sbagliate inaspettate, come ad esempio "20", "500" o "1". Per ogni CRT il rapporto di errori "impulsivi" su "altri" errori è più elevato per le femmine piuttosto che per i maschi. In questo modo, i dati ci suggeriscono che i maschi sono più propensi a riflettere sulle loro risposte e meno inclini a cadere nelle risposte impulsive.

In Appendice B è mostrata visivamente la differenza di genere nella prestazione "impulsiva" dei tre CRT. In una scala da 0 a 3 punti, in media, nel nostro esperimento è stato ottenuto un punteggio molto alto e pari a di 1,66, diviso in 1,51 in media per i maschi e 1,88 in media per le femmine. Anche per questa ragione il CRT funge da "trappola cognitiva" causata dal Sistema 1 piuttosto che dal Sistema 2. Anche Ponti e Cueva (2015) hanno generato una "iCRT" uguale al nostro "CRT Score Impulsivo" e in questo caso hanno osservato delle medie elevate e pari a 1,46 per i maschi e 1,93 per le femmine. Abbiamo testato tale differenza nelle prestazioni maschio/femmina usando il Mann-Whitney Test e abbiamo trovato un Effetto Genere statisticamente significativo con un p-value di 0.0548²¹.

²¹ Il t-test ha un p-value di 0.0294.

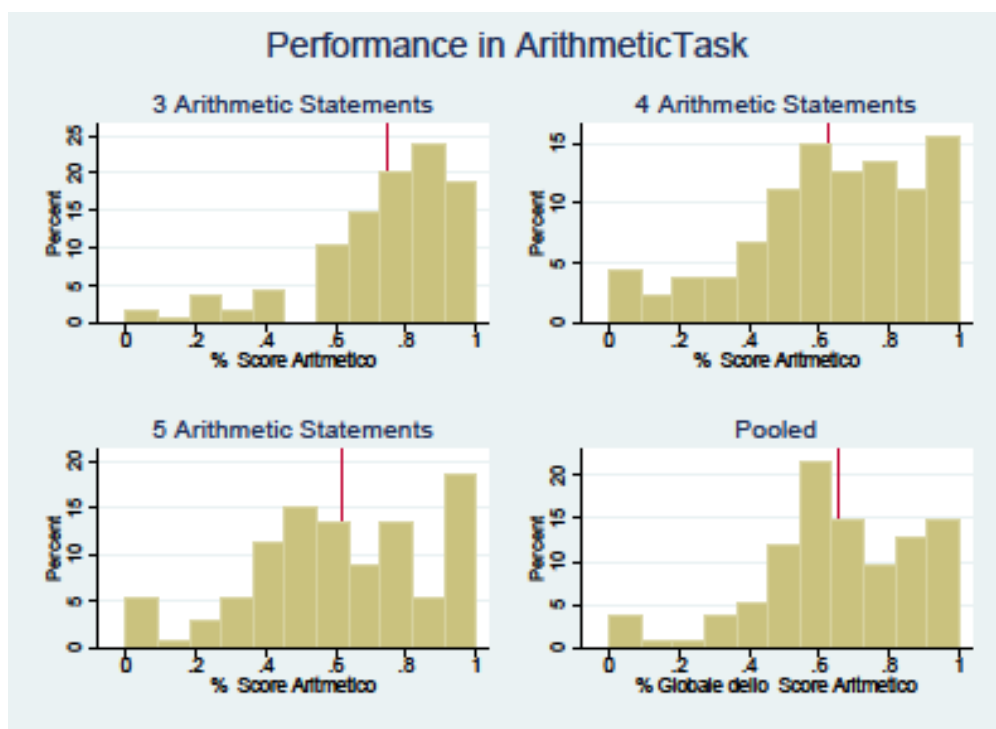
Risultati

In questa sezione mostriamo i nostri risultati generali.

Prima di tutto, verifichiamo la nostra **Ipotesi 3**: Il numero di parole ricordate (come proxy della WM a breve termine) si riduce con l'aumentare del numero delle parole da mantenere in memoria.

Controlliamo i risultati del nostro Operation Span Task. Esso si divide in due differenti tipi di attività da svolgere contemporaneamente. L'attività di interferenza (equazione aritmetica) e la proxy della Working Memory a breve termine (la parola da ricordare correttamente). Durante lo svolgimento di questo Task viene aumentato il numero di parole da ricordare e delle equazioni aritmetiche, per porre un maggior carico cognitivo sulla Working Memory a breve termine. Il Task comincia presentando, una alla volta, prima tre poi quattro e per ultimo cinque parole ed equazioni aritmetiche. Inizialmente, osserviamo la performance sulle equazioni aritmetiche. Le medie sono il 74,7% risolte quando le equazioni erano 3; il 62,9% risolte quando le equazioni erano 4; 61,7% risolte quando le equazioni erano 5; la media generale è stata di 65,3% di equazioni risolte. Abbiamo impostato dei t-test tra le performance di questi diversi trattamenti, al fine di verificare il livello di significatività di queste differenze nelle performance.

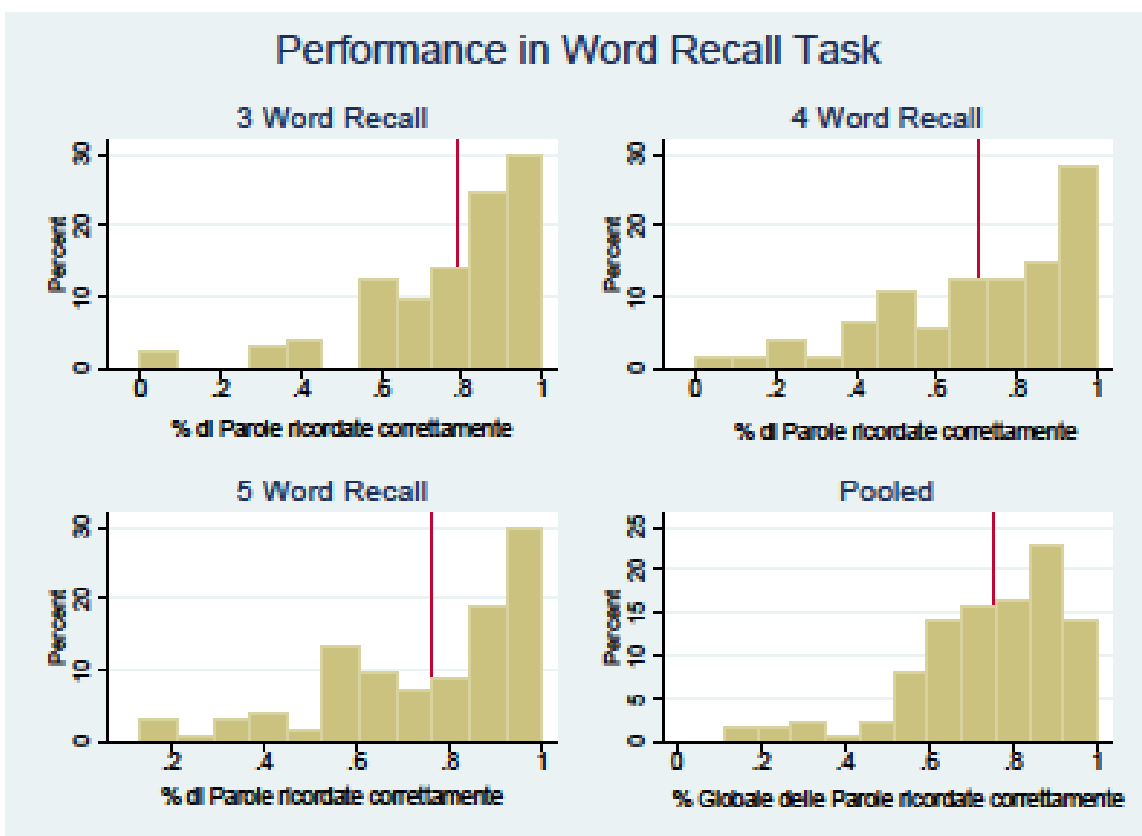
Grafico 2



Il Grafico 2 mostra la performance dei partecipanti sia nei singoli Task, sia quella complessiva. Confrontando i nostri risultati emerge che la performance di quando le equazioni erano 3 risulta maggiore di quando le stesse erano 4 e 5 (con i rispettivi p-value di 0,001 in entrambi i casi). Mentre, non si è rivelata nessuna differenza nelle performance statisticamente significativa tra quando le equazioni erano 4 con quando erano 5. Questo suggerisce che i partecipanti abbiano più difficoltà nel risolvere questi problemi quando ce ne sono un numero maggiore, ma questo può essere confutato da un possibile Effetto Ordine (i partecipanti potrebbero annoiarsi o stancarsi).

Successivamente, osserviamo la performance sulle parole da ricordare. Le medie sono il 78,7% ricordate quando le parole erano 3; il 70,6% ricordate quando le parole erano 4; il 76% ricordate quando le parole erano 5; la media generale è stata di 74,9% di parole ricordate. Abbiamo impostato dei t-test tra le performance di questi diversi trattamenti, al fine di verificare il livello di significatività di queste differenze nelle performance.

Grafico 3



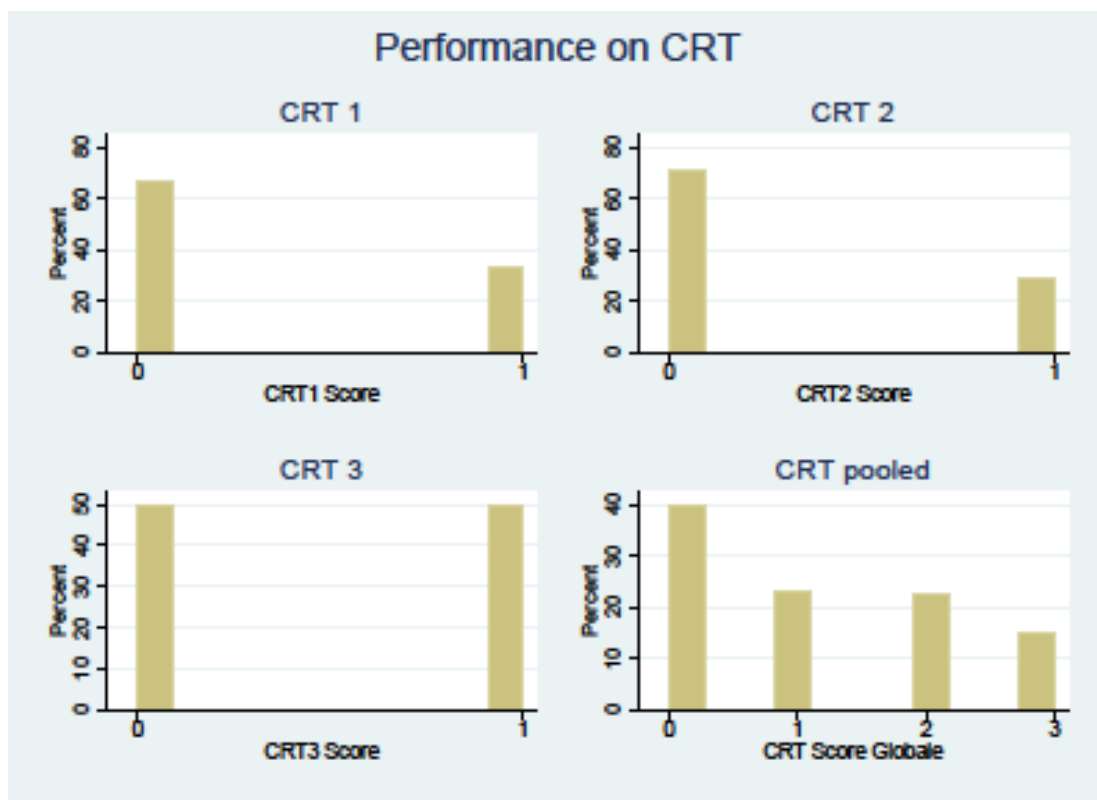
Il Grafico 3 mostra la performance dei partecipanti sia nei singoli Task, sia quella complessiva. Confrontando i nostri risultati emerge che la performance di quando le parole erano 3 risulta maggiore di quando le stesse erano 4 (p-value di 0.001). Non abbiamo osservato nessuna differenza nella performance di quando le parole da ricordare erano 3 rispetto a quando erano 5. Inoltre, la performance di quando le parole erano 5 si è rivelata maggiore confrontata con la performance di quando le stesse erano 4. Pertanto, la performance è maggiore quando le parole da ricordare sono 3 e 5. Questo suggerisce che i risultati sono influenzati dall' Effetto Ordine. Questo effetto si verifica quando l'ordine delle attività da svolgere è sempre lo stesso (3 parole da ricordare, poi 4 e infine 5), in quanto i partecipanti possono imparare dai primi trattamenti e di conseguenza ricordare più parole. L'Effetto Ordine per le equazioni aritmetiche è probabile che non sia apprendimento, in quanto i partecipanti nel tempo ottengono performance minori. Qui invece sembra che la performance su 4 parole sia minore confrontata con quella su 3 parole, mostrando un incremento di difficoltà. Il passaggio dei trattamenti da 4 a 5 parole sembra mostrare un incremento nelle performance, indicando apprendimento (Effetto Ordine). Per questa ragione, non possiamo confermare con certezza la nostra terza ipotesi.

Verifichiamo la nostra **Ipotesi 4**: Il numero di parole ricordate (come proxy della WM a breve termine) si riduce con l'aumentare della lunghezza delle parole da mantenere in memoria.

Possiamo controllare i risultati del nostro Operation Span Task, aggiungendo una variabile dummy che riflette la lunghezza della parola da ricordare. Abbiamo inoltre impostato dei t-test tra i trattamenti con differenti numero di parole da ricordare al fine di verificare il livello di significatività nelle differenze delle performance. Le medie percentuali globali sono state del 91,7% di parole "corte" ricordate e del 71% di parole "lunghe" ricordate (p-value di 0.001). Quando le parole erano 3, la differenza è stata minore, riscontrando un 86,1% di parole "corte" ricordate rispetto al 77% di parole "lunghe" ricordate (p-value di 0.0435). In aggiunta, quando le parole da ricordare erano 4 abbiamo notato una maggiore differenza nelle performance, riscontrando un 94,8% di parole corte ricordate e un 65% di parole lunghe ricordate (p-value di 0.001). In ultima analisi, quando le parole erano 5 abbiamo visto che la percentuale di parole corte ricordate era del 93% mentre per le parole lunghe era del 72% (p-value di 0.001). Pertanto, la nostra Ipotesi 4 è confermata, anche se è possibile notare un Effetto Ordine che influenza i risultati. Le performance sulle parole corte sembrano migliorare già nel passaggio dalle 3 alle 4 parole, mentre con le parole lunghe l'apprendimento è osservabile nel passaggio dalle 4 alle 5 parole (Effetto Ordine).

Successivamente, cominciamo ad analizzare le performance sui Cognitive Reflection Test. Generiamo una variabile dipendente costruita assegnando un punto ai partecipanti per ogni risposta corretta sui tre CRT. E' la nostra proxy che stabilisce l'efficiente utilizzo del processo cognitivo deliberativo. In una scala da 0 a 3 punti, in media, nel nostro esperimento è stato ottenuto un punteggio di 1,13 di risposte corrette nei tre CRT. Possiamo affermare che sia una media modesta, ma non è sorprendente in quanto anche Frederick (2005) ha osservato una media di risposte corrette nei tre CRT simile alla nostra e pari a 1,24. Frederick (2005) ha spiegato inoltre come l'oggetto dei tre CRT siano "facili" nel senso che le loro soluzioni sono di semplice comprensione una volta spiegate. Tuttavia, per raggiungere le risposte corrette è spesso richiesta la soppressione della risposta sbagliata che "impulsivamente" è stata richiamata dalla mente.

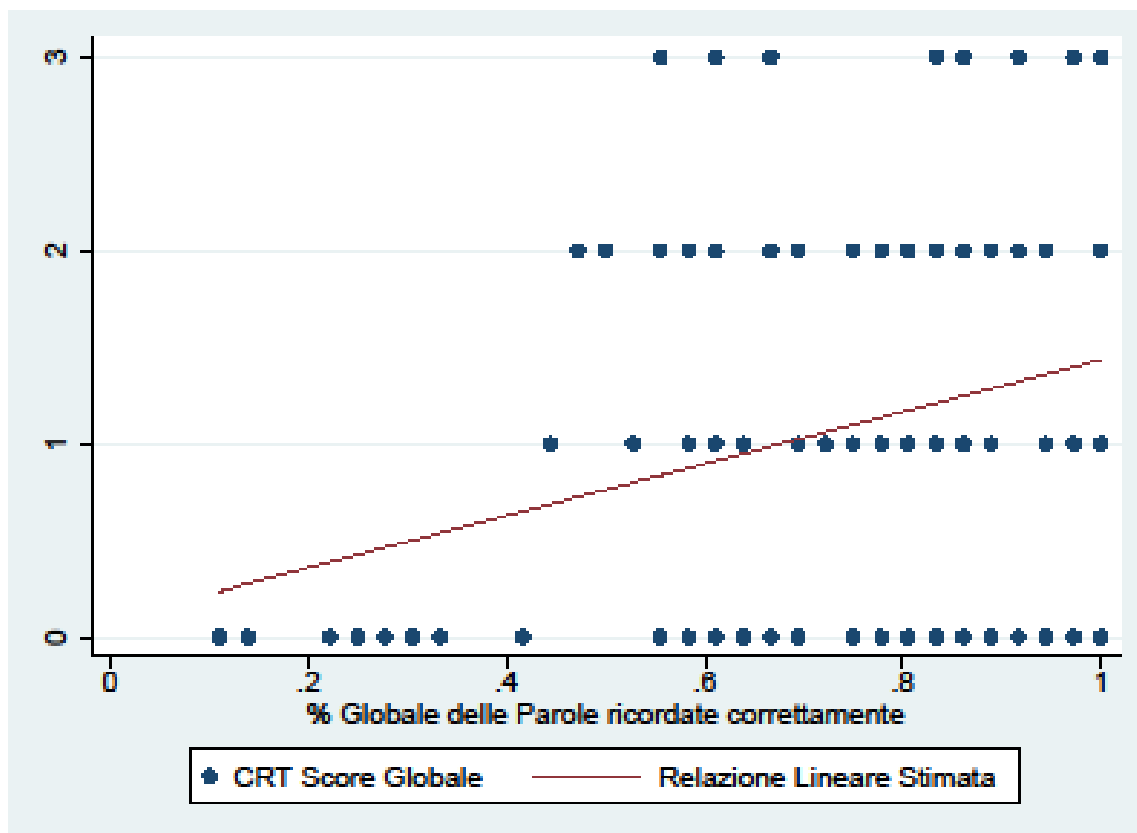
Grafico 4



Il Grafico 4 mostra sia la media percentuale di risposta corretta nei singoli CRT, sia la percentuale globale del numero di risposte corrette sui tre CRT. Come in letteratura, abbiamo osservato una media bassa di risposte corrette nonostante i CRT non richiedano particolari abilità matematiche e siano "facili" nel senso che sono facilmente comprensibili una volta spiegati. Le medie sono del 33,58% di probabilità di risposta esatta per il CRT 1, di 29,10% per

il CRT 2 e del 50% per il CRT 3. Facendo una analisi del CRT Score Globale otteniamo che solo il 14,93% dei partecipanti ha dato 3 risposte corrette, il 22,39% ha dato 2 risposte corrette, il 23,13% ha dato una risposta corretta e che il 39,55% ha dato neanche una risposta corretta. Per esempio, Garza, Kujal e Lenkei (2015) hanno trovato il 32% di probabilità di risposta corretta per il CRT 1, il 40% per il CRT 2 e il 48% per il CRT 3. Hanno anche ottenuto che solo il 18% dei partecipanti ha risposto correttamente a tutte le domande, il 21% ha risposto correttamente a due domande, il 22% dei partecipanti ha fornito una risposta corretta e il 38% non ne ha fornita alcuna corretta. Queste percentuali sono molto simili alle nostre tranne per quanto riguarda il CRT 2 in cui meno dei nostri partecipanti ha risposto correttamente. Anche Frederick (2005), nel suo lavoro ha mostrato delle medie percentuali simili alle nostre e pari a 17% per le tre risposte corrette, al 23% per due, al 28% per una e pari al 33% per nessuna risposta corretta.

Grafico 5. Relazione tra il CRT Score Globale e la Working Memory a breve termine

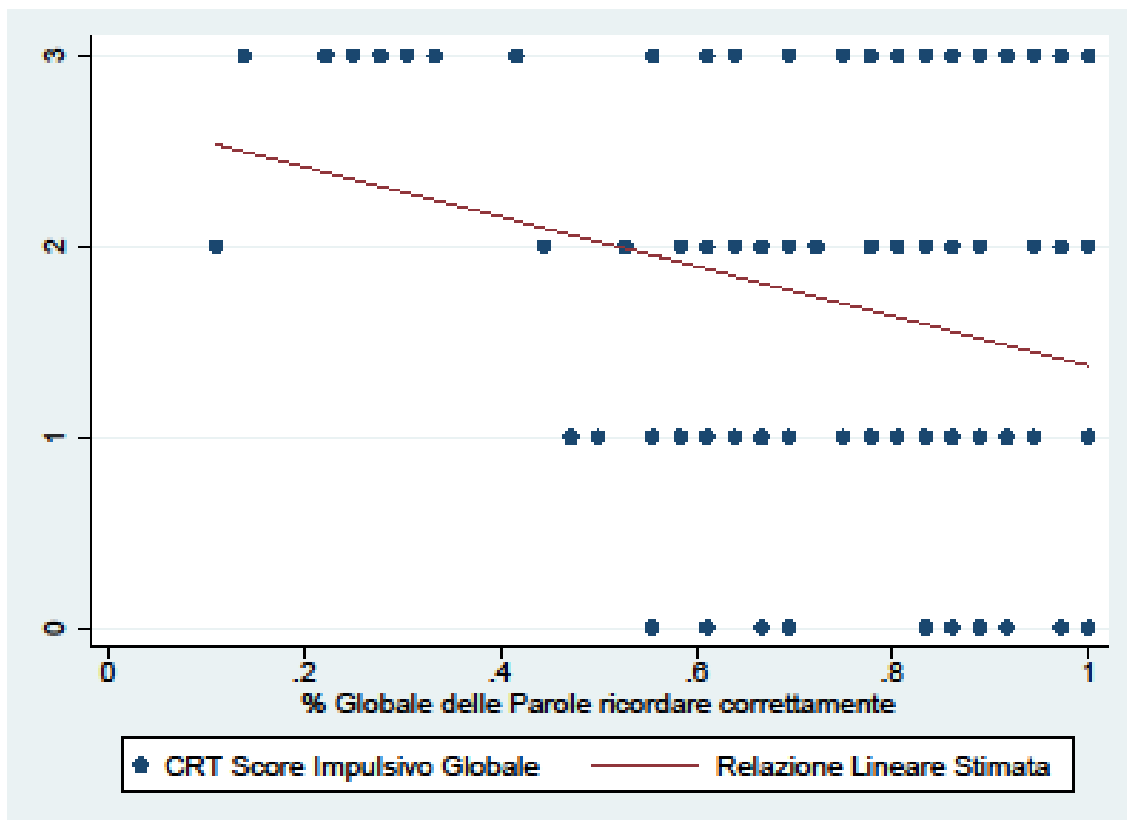


Il Grafico 5 mostra la relazione positiva tra lo score dei CRT basati sulla risposta corretta e la percentuale di parole ricordate correttamente. Questa sembra dare un leggero supporto alla nostra principale ipotesi, ovvero che la probabilità della corretta efficienza delle funzioni di

controllo del processo cognitivo deliberativo migliora con l'aumentare della nostra proxy della Working Memory a breve termine (la percentuale di parole ricordate correttamente).

Abbiamo esteso la nostra analisi usando un modo alternativo di analizzare i dati, attraverso una nuova variabile dipendente chiamata "CRT Score Impulsivo", illustrata precedentemente nelle statistiche descrittive. Questa variabile, speculare al precedente score, è costruita assegnando un punto ai partecipanti per ogni risposta impulsiva sui CRT. E' la nostra proxy che stabilisce il fallimento delle funzioni di monitoraggio del processo cognitivo deliberativo e la prevalenza del Sistema 1.

Grafico 6 Relazione tra il CRT Score Impulsivo Globale e la WM a breve termine



Il Grafico 6 mostra, invece, la relazione negativa tra il CTR Score Impulsivo basato sulla risposta impulsiva e la percentuale di parole ricordate correttamente (la nostra proxy per la Working Memory a breve termine). Questa relazione sembra dare un'ulteriore supporto alla nostra principale ipotesi, sotto un altro punto di vista. Il fallimento delle funzioni di monitoraggio del processo cognitivo deliberativo e la prevalenza del Sistema 1 sono incentivati da una minore Working Memory a breve termine (misurata dalla percentuale di parole ricordate

correttamente). In altre parole, lo Score Impulsivo sui CRT è inteso come la misura dell'incapacità di sopprimere le risposte impulsive erronee (Ponti, Cueva et al. 2015). In Appendice C possiamo trovare sia i valori medi percentuali di risposta impulsiva nei singoli CRT, sia la percentuale globale di risposte impulsive sui tre CRT. Tuttavia, si nota anche dal Grafico 6 che la retta, inclinata negativamente, è traslata verso l'alto rispetto a quando il CRT Score rifletteva le risposte corrette. In una scala da 0 a 3 punti, in media, nel nostro esperimento è stato ottenuto un punteggio di 1,66 di risposta impulsiva nei tre CRT. Abbiamo impostato dei t-test per verificare le differenze tra gli score "impulsivi" e "corretti". Nel CRT 1 in media, abbiamo ottenuto una prestazione maggiore nello score impulsivo rispetto a quello corretto con un p-value di 0,0001. Anche nel CRT 2 in media, abbiamo ottenuto lo stesso risultato con un p-value di 0,0001. Invece nel CRT 3 in media, non abbiamo ottenuto alcuna differenza statisticamente significativa tra gli score impulsivi e corretti. Mentre nel CRT Score Globale, abbiamo ottenuto nuovamente una prestazione maggiore nello score impulsivo rispetto a quello corretto con un p-value di 0.0024. Quindi in generale, sembra adeguato affermare che in media nei CRT è più alta la probabilità di rispondere impulsivamente piuttosto che utilizzare efficientemente le funzioni di monitoraggio del Sistema 2. Questa affermazione è coerente con la teoria di Frederick (2005) e i risultati di Ponti e Cueva (2015).

Entriamo nel dettaglio della nostra analisi e usiamo due step per verificare le nostre principali ipotesi.

Ipotesi 1: Le Performance nei CRT migliorano con la Working Memory a breve termine misurata attraverso le parole da ricordare.

Assumiamo che la percentuale di parole ricordate correttamente sia una proxy della Working Memory a breve termine. La nostra principale ipotesi congiunta stabilisce un ruolo cruciale della Working Memory a breve termine per la comprensione della variazione dello Score dei CRT. Una migliore WM a breve termine dovrebbe aumentare la probabilità dei partecipanti di utilizzare efficientemente le funzioni di monitoraggio del Sistema 2, dando la risposta corretta.

Ipotesi 2: L'Effetto Genere deteriora le performance delle femmine nei CRT.

Questa ipotesi stabilisce sia che le femmine hanno più probabilità di cedere alla prevalenza del Sistema 1 rispetto ai maschi, sia che i maschi hanno più probabilità di usare efficientemente le funzioni di monitoraggio del Sistema 2 rispetto alle femmine.

Il primo step consiste nel controllare le relazioni tra la percentuale di parole ricordate correttamente (% Parole Ricordate è la nostra proxy della Working Memory a breve termine), educazione e reddito dei genitori.

Tab. 5 Relazione tra la percentuale di parole ricordate, Educazione e Reddito dei genitori.

Correlazione	% Parole Ricordate	Reddito Genitori
% Parole Ricordate		
Reddito Genitori	12.08%	
Educazione Padre	9.69%	39.36%
Educazione Madre	-1.57%	18.52%

La Tabella 5 mostra che la correlazione tra % Parole Ricordate e Reddito è leggermente positiva e pari a 12,08%. Invece, guardando la correlazione tra % Parole Ricordate e Educazione dei genitori notiamo due effetti opposti. La nostra proxy della WM a breve termine è correlata positivamente con l'Educazione del Padre (9,69%) e negativamente con l'Educazione della Madre (-1,57%). Infine, la correlazione tra Reddito ed Educazione dei genitori è positiva. Questo potrebbe significare che il livello di educazione della madre dei partecipanti potrebbe influire negativamente sulla loro performance nel Complex Span Task (riferito alla percentuale di parole ricordate correttamente). Al contrario, l'Educazione del Padre sembra influire positivamente.

In questo step impostiamo una regressione usando come variabile dipendente la nostra proxy per la WM a breve termine (la percentuale di parole ricordate correttamente). Le variabili indipendenti sono il Reddito dei Genitori, l'Educazione del Padre, l'Educazione della Madre, Età, Genere (1 femmina ; 0 maschio), Major (1 corso di studi inerenti la matematica ; 0 altrimenti) e il GPA.

$$reg \% Recall = \beta_0 + \beta_1 Inc. of Parents + \beta_2 Ed. of Father + \beta_3 Ed. of Mother + \beta_4 Age + \beta_5 Gender + \beta_6 Major + \beta_7 GPA + e$$

Tuttavia, abbiamo trovato che tutte le variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%. Potrebbe significare che le postulate variabili indipendenti non abbiamo effetto sulla WM a breve termine coeteris paribus.

Il secondo step è diviso in tre parti:

1. Impostiamo due regressioni differenti focalizzandoci sul CRT Score, basato sulle risposte corrette, come variabile dipendente e cambiando il set di variabili indipendenti. Prima consideriamo nella regressione il Reddito dei Genitori, l'Educazione del Padre e della Madre, e successivamente le escludiamo.
2. Svolgiamo lo stesso procedimento, focalizzandoci sul CRT Score Impulsivo, basato sulle risposte impulsive, come variabile dipendente.
3. Mettiamo a confronto i due gruppi "estremi", nel senso che: uno riflette chi ha dato tutte e tre le risposte corrette, l'altro riflette chi ha dato tutte e tre le risposte impulsive.

I modelli di regressione logit sono modelli di regressione non-lineare che vengono utilizzati quando la variabile dipendente è una dummy. In questo caso, la variazione nella variabile dipendente per effetto della variazione di una delle indipendenti è pari alla variazione nella probabilità che la variabile dipendente sia uguale a uno. Le regressioni logit utilizzano la funzione di ripartizione logistica standard e i coefficienti sono stimati con la tecnica della massima verosimiglianza. In Appendice E è possibile osservare tutte le regressioni impostate su STATA.

Nella Parte 1, usiamo il CRT Score come proxy del corretto funzionamento del processo cognitivo deliberativo (il Sistema 2). In primo luogo, impostiamo una regressione logit per ogni CRT preso singolarmente. In secondo luogo, impostiamo la regressione logit sul CRT score

globale. In terzo luogo impostiamo le stesse regressioni escludendo le variabili riguardanti l'educazione e il reddito dei genitori.

Impostiamo tre regressioni logit sulla nostra variabile dipendente (1 risposta corretta; 0 altrimenti) per ogni CRT. Prendiamo in considerazione come variabili indipendenti: la percentuale di parole ricordate correttamente (come proxy della WM a breve termine), l'educazione del padre e della madre, il reddito dei genitori, il genere, l'età, il major e il GPA. La prima è la nostra variabile indipendente principale, mentre le altre sono le variabili demografiche.

$$\text{logit CRT } i \text{ Score} = \beta_0 i + \beta_1 \% \text{ Recall} + \beta_2 \text{ Inc. of Parents} + \beta_3 \text{ Ed. of Father} \\ + \beta_4 \text{ Ed. of Mother} + \beta_5 \text{ Age} + \beta_6 \text{ Gender} + \beta_7 \text{ Major} + \beta_8 \text{ GPA} + e_i$$

$i = 1, 2, 3$

Nella regressione sul CRT 1 Score, abbiamo trovato che la percentuale di parole ricordate correttamente ha un beta positivo e un p-value di 0,09; gender ha un coefficiente negativo e un p-value di 0,044; major ha un beta positivo e un p-value di 0,035; e GPA ha un coefficiente positivo e un p-value di 0,074. Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Nella regressione sul CRT 2 Score, abbiamo trovato che la percentuale di parole ricordate correttamente ha un beta positivo e un p-value di 0,019; educazione del padre ha un coefficiente positivo ed un p-value di 0,058; educazione della madre ha un beta negativo e un p-value di 0,070; età ha un coefficiente positivo e un p-value di 0,003; gender ha un beta negativo e un p-value di 0,068; e GPA ha un coefficiente positivo e un p-value di 0,041. Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Nella regressione sul CRT 3 Score, abbiamo trovato che la nostra principale variabile indipendente ha un beta positivo e un p-value di 0,007; età ha un coefficiente positivo e un p-value di 0,074; gender ha un beta negativo e un p-value di 0,086; e GPA ha un coefficiente positivo e un p-value di 0,022. Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Chiaramente per il CRT Score Globale, abbiamo impostato la regressione logit multinomiale in quanto può assumere 4 valori (0, 1, 2, 3).

$$m \logit \text{ Global CRT Score} = \beta_0 + \beta_1 \% \text{ Recall} + \beta_2 \text{ Inc. of Parents} + \beta_3 \text{ Ed. of Father} + \beta_4 \text{ Ed. of Mother} + \beta_5 \text{ Age} + \beta_6 \text{ Gender} + \beta_7 \text{ Major} + \beta_8 \text{ GPA} + e$$

La nostra principale variabile indipendente ha un beta positivo e un p-value sempre basso (0.010, 0.028, 0.025 per i valori rispettivamente di 1, 2, 3 dei CRT Score); gender ha sempre un coefficiente negativo e un basso p-value (10% o inferiore); e GPA ha sempre un coefficiente positivo e un p-value basso. Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Come possiamo vedere, c'è un filo che lega tutte queste regressioni quando sono inclusi l'educazione del padre e della madre e i loro redditi. La nostra principale variabile indipendente è fondamentale per comprendere il cambiamento degli Score nei CRT. La percentuale di parole ricordate ha sempre un coefficiente positivo e un basso p-value (10% o inferiore). Pertanto, la nostra prima ipotesi sembra confermata, le performance sui CRT migliorano con la WM a breve termine misurata attraverso le parole ricordate. Il Reddito dei Genitori è sempre non statisticamente significativo, in quanto una sua variazione non implica una variazione di probabilità della nostra variabile dipendente. Il GPA ha sempre un beta positivo e un p-value vicino allo 0.05, in altre parole un aumento del GPA permette una più probabile attivazione del Sistema 2. L'Effetto Genere è sempre presente e ci mostra come i maschi usino maggiormente il processo cognitivo deliberativo rispetto alle femmine, confermando così la nostra seconda ipotesi. Major ha un beta positivo ed è statisticamente significativo solo nel CRT1, questo sembra avere un debole effetto positivo sullo score dei CRT. L'Età è statisticamente significativa solo nei CRT 2 e 3 e sembra avere un effetto positivo sullo score dei CRT. In altre parole, l'aumento dell'età sembra portare un leggero beneficio in termini di probabilità di efficiente utilizzo del Sistema 2. C'è una sorpresa nella regressione del CRT 2, in questo caso l'educazione del padre è una variabile consistente e ha un beta positivo, anche l'educazione della madre è statisticamente significativa ma ha un coefficiente negativo. Questo significa che lo score del CRT 2 aumenta se è maggiore l'educazione del padre o è minore l'educazione della madre, ma questi effetti sembrano deboli. Nelle altre regressioni, queste ultime variabili non sono statisticamente significative a un livello di confidenza del 10%.

Impostiamo le stesse regressioni precedenti escludendo il Reddito dei genitori e l'Educazione del Padre e della Madre dei partecipanti.

$$\text{logit CRT } i \text{ Score} = \beta_0 + \beta_1 \% \text{ Recall} + \beta_2 \text{ Age} + \beta_3 \text{ Gender} + \beta_4 \text{ Major} + \beta_5 \text{ GPA} + e_i$$

$i = 1, 2, 3, \text{ global}$

I risultati sono gli stessi, in altre parole la nostra proxy della WM a breve termine è statisticamente significativa (p-value del 10% o inferiore) e ha un beta positivo, migliorando così lo score dei CRT. L' Effetto Genere nei CRT è consistente e ha un coefficiente negativo che conferisce un vantaggio per i maschi. Il GPA è consistente ed ha un effetto positivo sullo Score dei CRT. L'età è statisticamente significativa ed ha un effetto positivo solo nello score del CRT 1 e 2. Il Major è statisticamente significativo e ha un effetto positivo solo nel CRT 1, in quanto mostra un debole effetto.

Nella Parte 2, usiamo il CRT Score Impulsivo come proxy per indicare il fallimento nella funzione di monitoraggio del Sistema 2 e conseguente prevalenza del processo cognitivo euristico (Sistema 1). In primo luogo, impostiamo una regressione logit per ogni CRT preso singolarmente. In secondo luogo, impostiamo la regressione logit sul CRT Score Impulsivo Globale. In terzo luogo impostiamo le stesse regressioni escludendo le variabili riguardanti l'educazione e il reddito dei genitori.

Impostiamo tre regressioni logit sulla nostra variabile dipendente (1 risposta impulsiva; 0 altrimenti) per ogni CRT. Prendiamo in considerazione come variabili indipendenti: la percentuale di parole ricordate correttamente (come proxy della WM a breve termine), l'educazione del padre e della madre, il reddito dei genitori, il genere, l'età, il major e il GPA. La prima è la nostra variabile indipendente principale, mentre le altre sono le variabili demografiche.

$$\begin{aligned} \text{logit CRT } i \text{ Score Impulsivo} \\ = \beta_0 + \beta_1 \% \text{ Recall} + \beta_2 \text{ Inc. of Parents} + \beta_3 \text{ Ed. of Father} \\ + \beta_4 \text{ Ed. of Mother} + \beta_5 \text{ Age} + \beta_6 \text{ Gender} + \beta_7 \text{ Major} + \beta_8 \text{ GPA} + e_i \end{aligned}$$

$i = 1, 2, 3$

Nella regressione sul CRT 1 Score Impulsivo, abbiamo trovato che la percentuale di parole ricordate correttamente non è statisticamente significativa al livello di confidenza del 10%, in quanto ha un p-value di 0,112; major ha un beta negativo e un p-value di 0,034; genere ha un beta positivo e un p-value di 0,042; e GPA ha un beta negativo e un p-value di 0,036. Tutte le

altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Nella regressione sul CRT 2 Score Impulsivo, abbiamo trovato che la percentuale di parole ricordate correttamente ha un beta negativo e un p-value di 0,052; educazione del padre ha un coefficiente negativo ed un p-value di 0,003; età ha un beta negativo e un p-value di 0,004; genere ha un coefficiente positivo e un p-value di 0,006; e GPA ha un coefficiente negativo e un p-value di 0,025. Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Nella regressione sul CRT 3 Score, abbiamo trovato che la nostra principale variabile indipendente ha un beta negativo e un p-value di 0,001; educazione della madre ha un beta positivo e un p-value di 0,019; età ha un coefficiente negativo e un p-value di 0,007; genere ha un beta positivo e un p-value di 0,096; e GPA ha un coefficiente negativo e un p-value di 0,005. Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Chiaramente per il CRT Score Impulsivo Globale, abbiamo impostato la regressione logit multinomiale in quanto può assumere 4 valori (0, 1, 2, 3), come nel caso multinomiale precedente.

$$\begin{aligned} m \text{ logit global Impulsive CRT score} \\ = \beta_0 + \beta_1 \% \text{ Recall} + \beta_2 \text{ Inc. of Parents} + \beta_3 \text{ Ed. of Father} \\ + \beta_4 \text{ Ed. of Mother} + \beta_5 \text{ Age} + \beta_6 \text{ Gender} + \beta_7 \text{ Major} + \beta_8 \text{ GPA} + e \end{aligned}$$

La nostra principale variabile indipendente ha un beta negativo e un p-value sempre basso (0.096, 0.089, 0.006 per i valori rispettivamente di 1, 2, 3 del CRT Score Impulsivo); educazione del padre ha sempre un beta negativo e un p-value basso (10% o inferiore); genere ha sempre un coefficiente positivo e un basso p-value (10% o inferiore); e GPA ha sempre un coefficiente negativo e un basso p-value (10% o inferiore). Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Come possiamo vedere, la percentuale delle parole ricordate correttamente (la nostra proxy della WM a breve termine), in questo caso, ha sempre un coefficiente negativo e un basso p-value (10% o inferiore). In altre parole, una migliore Working Memory a breve termine significa avere una minore probabilità di essere impulsivo, confermando i nostri risultati precedenti. Il Reddito dei Genitori è sempre non statisticamente significativo, come nelle regressioni precedenti. Il GPA ha sempre un beta negativo e un p-value vicino a 0,05, il quale

ci indica che un suo aumento si traduce in una diminuzione della probabilità di essere impulsivo. L'Effetto Genere è sempre presente, in particolare verifica una maggiore impulsività per le femmine rispetto ai maschi. Il Major è statisticamente significativo solo nel CRT 1 ed ha un beta negativo, confermando i nostri precedenti risultati. L'Età è statisticamente significativa solo nel CRT 2 e 3 e sembra avere un effetto negativo sull'impulsività. In altre parole una minore età porta ad avere una maggiore probabilità di essere impulsivo. Ci sono delle sorprese nelle regressioni sullo Score Impulsivo, qui la variabile Educazione del Padre è spesso statisticamente significativa (con un p-value del 10% o inferiore) ed ha un beta negativo, mentre la variabile Educazione della Madre è statisticamente significativa (con un p-value del 5% or less) ed ha un coefficiente positivo solamente nel Score Impulsivo del CRT 2 e del CRT 3. Questo può significare che la probabilità di essere impulsivo sembra diminuire con l'aumento del livello di educazione del padre dei partecipanti, mentre la stessa probabilità aumenta leggermente con l'aumento del livello di educazione della madre dei partecipanti.

Impostiamo le stesse regressioni precedenti escludendo il Reddito dei genitori e l'Educazione del Padre e della Madre dei partecipanti.

logit CRT i Score Impulsivo

$$= \beta_0 + \beta_1 \% Recall + \beta_2 Age + \beta_3 Gender + \beta_4 Major + \beta_5 GPA + e_i$$

i = 1, 2, 3, global

I risultati sono gli stessi, in altre parole la nostra proxy della WM a breve termine è statisticamente significativa (p-value del 10% o inferiore) e ha un beta negativo, peggiorando così lo Score Impulsivo dei CRT. L'Età è statisticamente significativa solo nel CRT 2 e nel 3 ed ha un effetto negativo sullo Score Impulsivo di entrambi. L'Effetto Genere nei CRT è consistente e ha un coefficiente negativo che conferisce un vantaggio per i maschi. Il Major è statisticamente significativo e ha un effetto negativo solo nel CRT 1, in quanto mostra un debole effetto. Il GPA è consistente ed ha un effetto negativo sullo Score Impulsivo dei CRT, in altre parole una diminuzione del GPA si traduce in una maggiore probabilità di essere impulsivi. L'unico cambiamento che verificiamo è che la nostra variabile proxy della WM a breve termine della regressione sullo Score Impulsivo del CRT 1 diventa statisticamente significativa con un p-value di 0,090 e il coefficiente resta negativo.

Nella Parte 3 per completare l'analisi, impostiamo due addizionali regressioni logit usando come variabile dipendente due diverse dummy che dividono i partecipanti in quattro gruppi "estremi".

La prima variabile dipendente assume valore "1" se i partecipanti danno 3 risposte corrette in tutti e tre i CRT e assume valore "0" altrimenti. Questo è un caso estremo in quanto divide i partecipanti in due gruppi, uno in cui i partecipanti utilizzano efficientemente la funzione di monitoraggio del Sistema 2 in tutti e tre i CRT, mentre nell'altro questo non si verifica. Nel nostro esperimento, questa variabile assume valore 1 solamente 20 volte e sospettiamo siano effettivamente poche osservazioni. Infatti, la nostra principale variabile indipendente non è statisticamente significativa ed ha un p-value di 0,106; e genere ha un coefficiente negativo e un p-value di 0,057. Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Se escludiamo il reddito e l'educazione dei genitori, l'unico cambiamento che verificiamo è che la percentuale di parole ricordate diventa significativa con un p-value di 0,076 e ha un coefficiente positivo, confermando i nostri risultati precedenti.

Per incrementare il numero delle osservazioni, noi generiamo anche altre due variabili dummy come variabili dipendenti. Una assume valore "1" se i partecipanti danno almeno due risposte corrette nei tre CRT e "0" altrimenti. L'altra assume valore "1" se i partecipanti danno almeno una risposta corretta nei tre CRT e "0" altrimenti. Tuttavia, queste regressioni non aggiungono niente di nuovo rispetto alle precedenti regressioni. In altre parole, confermiamo che la nostra proxy della Working Memory a breve termine è statisticamente significativa (con un p-value del 5% o inferiore) ed ha un coefficiente positivo. Queste ultime regressioni quindi, non danno maggiore chiarezza sulla relazione tra lo score dei CRT, l'educazione e il reddito dei genitori.

L'altra variabile dipendente assume valore "1" se i partecipanti danno le tre risposte impulsive nei tre CRT e "0" altrimenti. Questo è un caso estremo in quanto divide i partecipanti in due gruppi, uno in cui i partecipanti sono totalmente impulsivi in tutti e tre i CRT, mentre nell'altro questo non si verifica. Nel nostro esperimento questo è accaduto spesso, in quanto quasi un terzo dei soggetti si è comportato in questo modo (più del doppio dei partecipanti del caso speculare precedente). La nostra principale variabile indipendente ha un beta negativo e un p-value di 0,001; reddito dei genitori ha un coefficiente positivo e un p-value di 0,042; educazione del padre ha un beta negativo e un p-value di 0,093; educazione della madre non è

statisticamente significativa (p-value di 0,104); genere ha un coefficiente positivo e un p-value di 0,008; e GPA ha un coefficiente negativo e un p-value di 0,001. Tutte le altre variabili indipendenti non sono statisticamente significative a un livello di confidenza del 10%.

Questa è una regressione molto importante, in quanto ci fornisce numerose informazioni. L'aumento del valore della nostra proxy della WM a breve termine si traduce in una minore probabilità di essere impulsivi, confermando i nostri risultati precedenti. Il Reddito dei Genitori ha un effetto positivo e statisticamente significativo nel gruppo estremo delle tre risposte impulsive. Questo potrebbe significare che gli studenti provenienti da famiglie ricche sono più impulsivi, mentre quando il CRT Score è rappresentato dalla risposta corretta la variabile Reddito dei Genitori si rivelava statisticamente non significativa. L'Educazione del Padre è significativa ed ha sempre un coefficiente negativo, mentre l'educazione della madre ha un debole effetto sulla probabilità di dare la risposta impulsiva (p-value di 0,104). Rimane confermato l'Effetto Genere, il quale stabilisce una maggiore probabilità di dare le risposte impulsive per le femmine piuttosto che per i maschi. Anche l'effetto del GPA rimane confermato, un migliore GPA si traduce in una minore probabilità di essere impulsivo.

Per essere coerenti, abbiamo generato altre due variabili dummy come variabili dipendenti. Una assume valore "1" se i partecipanti danno almeno due risposte impulsive nei tre CRT e "0" altrimenti. L'altra assume valore "1" se i partecipanti danno almeno una risposta impulsiva nei tre CRT e "0" altrimenti. Tuttavia, queste regressioni non aggiungono niente di nuovo rispetto ai precedenti risultati. In altre parole, la nostra proxy della WM a breve termine è sempre statisticamente significativa (con un p-value del 10% o inferiore) ed ha un beta negativo. Quando le risposte impulsive sono almeno due, troviamo che l'educazione del padre è statisticamente significativa (p-value del 0,013) ed ha un coefficiente negativo, mentre l'educazione della madre è statisticamente significativa (p-value del 0,069) ma ha un beta positivo confermando i nostri risultati precedenti. Quando la risposta impulsiva è almeno una, troviamo che solo l'educazione del padre è statisticamente significativa (p-value di 0,046) ed ha un beta negativo. Questo potrebbe significare che l'educazione della madre ha un più debole effetto rispetto all'educazione del padre, oltre ad influenzare lo Score Impulsivo dei CRT in modo positivo. L'effetto negativo sulla probabilità di dare le risposte impulsive del GPA resta confermato (p-value del 10% o inferiore).

In ultima analisi, possiamo osservare che gli incentivi monetari permettono un leggero miglioramento nelle performance dei CRT. Abbiamo posto i CRT prima della conclusione dell'esperimento. La nostra media generale di risposte corrette sui tre CRT è stata molto bassa

e pari a 1,13. Sulla stessa linea con gli incentivi monetari, Frederick (2005) ha osservato una media simile alla nostra e pari a 1,24. Un altro esempio con incentivi reali, Oechssler et al. (2009) hanno osservato una media di 1,73. Hoppe e Kusterer (2011) anche hanno analizzato i CRT attraverso l'utilizzo degli incentivi reali per le prestazioni ed hanno osservato una media di 1,82. Dall'altra parte, senza incentivi monetari, Ponti e Cueva (2015) hanno osservato una media di 0,85. Un altro esempio senza incentivi reali, Bosch e Domenech (2014) hanno osservato una media di 0,75. Ancora un esempio è dato dal lavoro di Altmann, Falk e Grunewald (2015) che hanno verificato una media di 0,76 senza uno specifico pagamento per le risposte corrette nei CRT. Garza, Kujal e Lenkei (2015) hanno verificato che ponendo i CRT alla fine dell'esperimento, si abbassa la percentuale di ottenere una risposta corretta. Questo risultato è importante in quanto fornisce un sostegno indiretto alla tesi che l'attenzione può diminuire con il tempo e tale caratteristica è importante nei test cognitivi.

Conclusioni

Il CRT è diventato sempre più popolare e il più usato nel predire la riflessione in esperimenti economici e in psicologia. Troviamo una media modesta di risposte corrette nei tre CRT, confermando la loro funzione di “trappola” cognitiva causata dalla prevalenza del Sistema 1 piuttosto che del Sistema 2.

La nostra principale ipotesi congiunta è stata analizzata nello studio di Barret, Tugade e Engle (2004) e stabilisce che la Working Memory a breve termine (misurata dalla percentuale di parole correttamente ricordate) è correlata con l’attenzione controllata che guida i partecipanti verso l’utilizzo del processo cognitivo deliberativo il quale migliora le performance nei CRT.

In prima analisi, confermiamo la nostra principale ipotesi ovvero che la maggiore percentuale di parole ricordate, come proxy della Working Memory a breve termine, migliora le performance sui CRT. Nel primo step, verificiamo relazioni positive tra la percentuale di parole ricordate e l’educazione del padre e il reddito dei genitori, mentre tra la nostra proxy e l’educazione della madre evidenziamo una relazione negativa. Nel secondo step, regrediamo il CRT Score utilizzando come variabili indipendenti la percentuale di parole ricordate (come proxy della WM a breve termine), il genere, il GPA, il reddito dei genitori, l’educazione del padre e della madre, il major (che riflette le abilità matematiche dei partecipanti) e l’età. Il CRT Score la nostra proxy che stabilisce l’efficiente utilizzo del processo cognitivo deliberativo, in quanto costruita assegnando un punto ai partecipanti per ogni risposta corretta nei tre CRT. La nostra principale variabile indipendente è fondamentale per comprendere il cambiamento degli Score nei CRT. In particolare quando il CRT score riflette le risposte corrette, la performance sui CRT aumenta con una migliore Working Memory a breve termine (misurata attraverso la percentuale di parole ricordate correttamente). Questo è coerente con la teoria di Barrett, Tugade e Engle (2004), in cui il punteggio più alto nella WM a breve termine permette un controllo maggiore dell’attenzione il quale attiva le funzioni di monitoraggio del processo cognitivo deliberativo (Sistema 2). Per completare i risultati nel secondo step, estendiamo l’analisi usando un modo alternativo di lavorare i dati attraverso una nuova variabile dipendente chiamata “CTR Score Impulsivo”. Tra tutte le possibili risposte sbagliate che la gente possa dare, le risposte impulsive (10, 100 e 24) dominano. In altre parole, lo Score Impulsivo sui CRT è inteso come la misura dell’incapacità di sopprimere le risposte impulsive erronee (Ponti, Cueva et al. 2015). E’ la nostra proxy che stabilisce il fallimento delle funzioni di monitoraggio

del processo cognitivo deliberativo e la prevalenza del Sistema 1, in quanto costruita assegnando un punto ai partecipanti per ogni risposta impulsiva nei tre CRT. I nostri risultati dimostrano che, in generale, la probabilità di dare la risposta corretta sembra più bassa della probabilità di dare la risposta impulsiva, confermando la teoria di Frederick (2005) e i risultati di Ponti e Cueva (2015). Come ulteriore supporto alla nostra ipotesi principale, osserviamo che la migliore WM a breve termine riduce la probabilità di dare la risposta impulsiva.

In seconda analisi, abbiamo trovato che l'Effetto Genere definito per la prima volta da Frederick (2005) nei CRT è solido. Quando il CRT Score riflette le risposte corrette, i maschi ottengono una performance particolarmente migliore rispetto alle femmine. Quando, invece, lo Score riflette le risposte impulsive, sono le femmine ad ottenere un punteggio più alto. Se un ricercatore è interessato ai risultati generali sui CRT, dovrebbe fare attenzione e dividere le prestazioni tra maschi e femmine. Per esempio, se si usa il criterio delle tre risposte impulsive, allora la performance è sproporzionatamente sbilanciata verso le femmine. Nel senso che, le femmine sembrano essere più impulsive, mentre i maschi probabilmente usano più spesso il processo cognitivo deliberativo quando prendono una decisione.

In terza analisi, non possiamo dimostrare l'ipotesi di Turner e Engle (1989), in cui la percentuale di parole ricordate si deteriora con l'aumentare del numero di parole tenute in mente. In particolare, l'Effetto Ordine per le equazioni aritmetiche è probabile che non si riferisca all'apprendimento, in quanto i partecipanti nel tempo ottengono performance minori. Questo suggerisce che i partecipanti abbiano più difficoltà nel risolvere tali problemi quando ce ne sono un numero maggiore, ma questa affermazione può essere confutata da un possibile Effetto Ordine (i partecipanti potrebbero annoiarsi o stancarsi). Per quanto riguarda la nostra proxy della WM a breve termine, invece, sembra che la performance su 4 parole sia minore confrontata con quella su 3 parole, mostrando un incremento di difficoltà. Il passaggio dei trattamenti da 4 a 5 parole sembra mostrare un incremento nelle performance, indicando apprendimento (Effetto Ordine). Per questa ragione, non possiamo confermare con certezza la nostra terza ipotesi.

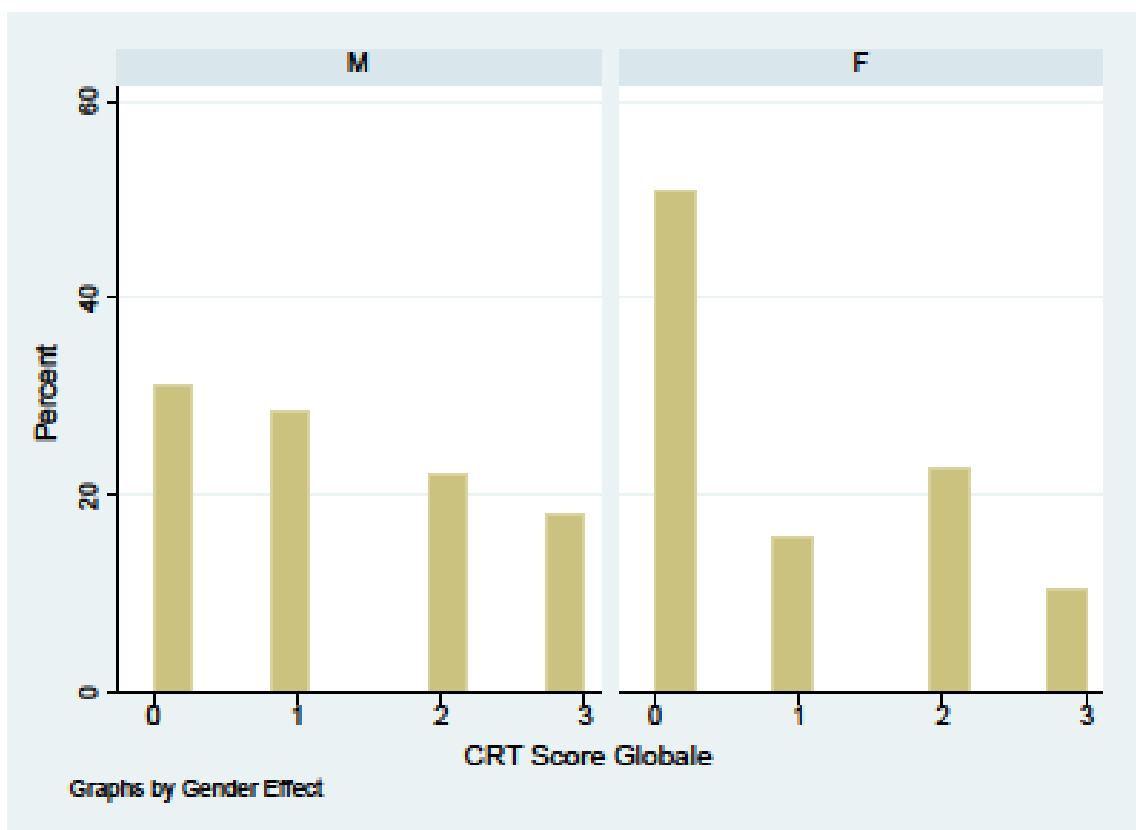
In quarta analisi, osserviamo che la percentuale di parole ricordate si deteriora con l'aumentare della lunghezza delle parole tenute in mente. In quanto i partecipanti ricordano una maggiore percentuale di parole "corte" rispetto a quelle "lunghe" (Baddeley, Thomson, e Buchanan, 1975).

In conclusione, cosa rimane dopo la lettura di questa tesi è un ulteriore supporto alle teorie del dual-process. Nonostante le altre teorie economiche si focalizzino esclusivamente sul

comportamento ottimale del pensiero razionale, la DPT conferma l'esistenza di due processi cognitivi nel pensiero razionale che competono tra loro per aiutarci a prendere le decisioni. Uno è un processo euristico, in cui non utilizziamo in modo ottimale le informazioni disponibili e siamo esposti ai bias cognitivi che limitano il nostro comportamento razionale. L'altro è un processo deliberativo, le informazioni sono analizzate efficientemente e il comportamento è completamente razionale. Il primo fattore chiave nella prevalenza dei due processi cognitivi è la Working Memory a breve termine che è positivamente correlata con la probabilità di usare il processo deliberativo per prendere decisioni. Il secondo fattore chiave è l'Effetto Genere, in cui le femmine hanno più probabilità di essere impulsive (bias cognitivo) e i maschi di usare la via completamente razionale. Tuttavia, i risultati osservati negli score dei CRT rimangono suscettibili di interpretazioni. La struttura causale può essere che il reddito e l'educazione dei genitori aumentino la capacità dei loro figli ad usare l'attenzione e a migliorare l'efficienza del processo cognitivo deliberativo quando necessario. Per questo motivo, queste variabili dovrebbero causare una Working Memory a breve termine migliore. Tuttavia, questa affermazione può invertire la causalità in modo tale da arrivare a conclusioni sbagliate. La causalità potrebbe essere che sia la WM a breve termine e sia altri fattori che dipendono dal reddito e dall'istruzione dei genitori concorrono a determinare le prestazioni sui CRT. Alcuni test per gli effetti demografici sono stati fatti spesso, ma una teoria per la quale andrebbero analizzati sembra mancare. In assenza di una tale teoria non è del tutto chiaro cosa esprime un effetto statisticamente non significativo. Sugeriamo che gli effetti demografici siano un altro fattore importante per comprendere meglio l'interazione tra la Working Memory a breve termine e le altre caratteristiche cognitive, emotive e culturali dell'individuo. Nell'immediato futuro ci proponiamo di continuare la stessa linea di ricerca incrementando la popolazione campionaria e la sua eterogeneità, al fine di ottenere una più ampia varietà di effetti demografici e di abilità cognitive.

APPENDICE

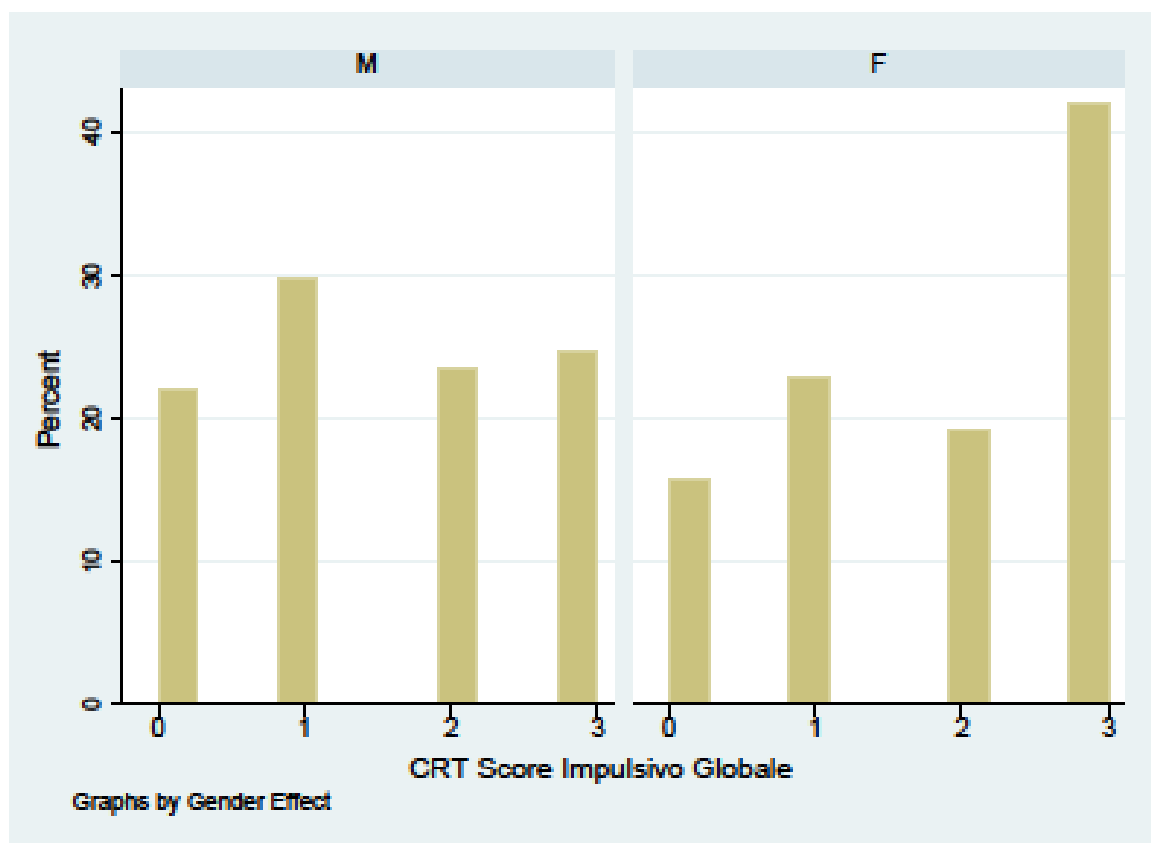
Appendice A Performance on Global CRT Score by Gender Effect



L'Appendice A mostra le performance nel CRT Score Globale, basato sulle risposte corrette, diviso per genere. Come si può notare la probabilità di zero risposte corrette è quasi il doppio per le femmine rispetto che per i maschi.

Appendice B

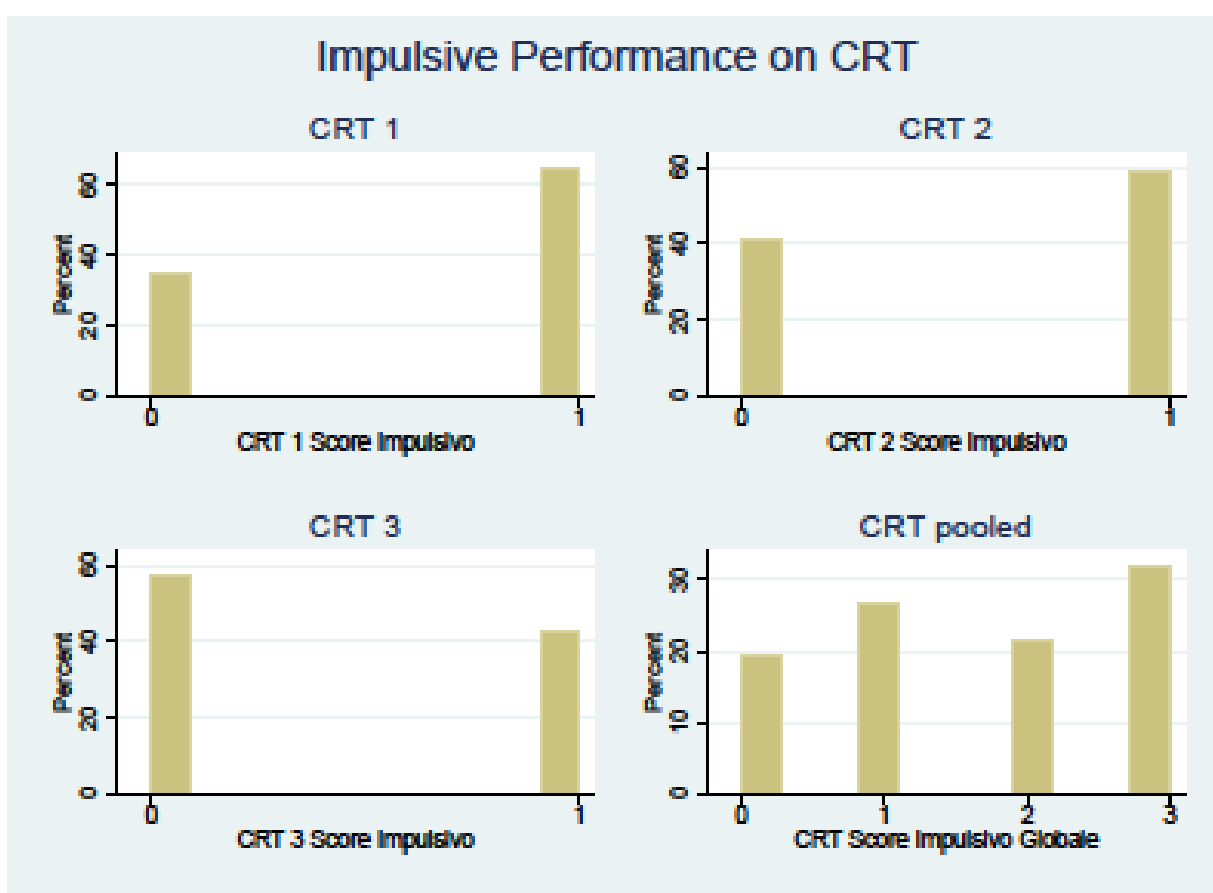
Performance on Impulsive Global CRT Score by Gender Effect



L'Appendice B mostra le performance nel CRT Score Impulsivo Globale, quindi basato sulle risposte impulsive, diviso per genere. Come si può notare la probabilità di tre risposte impulsive è quasi il doppio per le femmine rispetto che per i maschi.

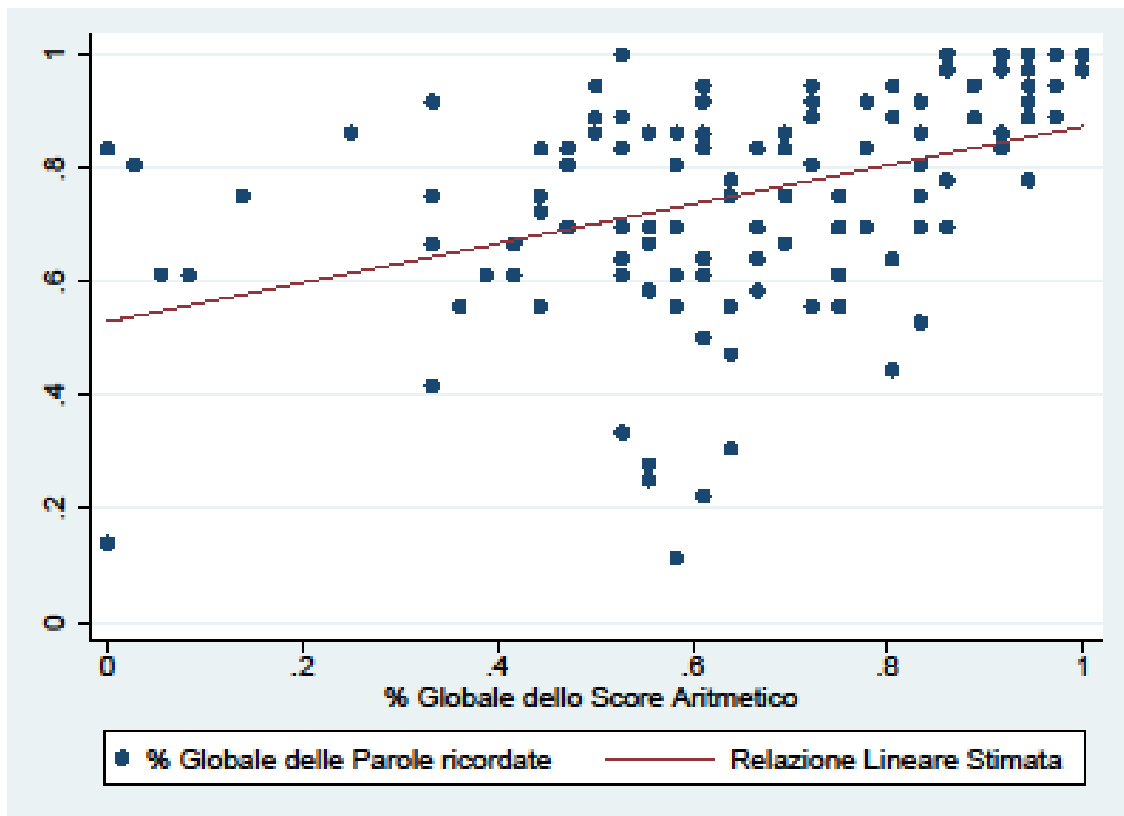
Appendice C

Impulsive Performance on CRT



L'Appendice C ci permette di evidenziare che, in generale, la percentuale di risposte impulsive sembra maggiore della percentuale di risposte corrette (opportuni t-test nella sezione dei risultati).

Appendice D Relationship between % of words to recall and % of arithmetic score



L'Appendice D mostra la relazione positiva tra la percentuale di parole ricordate correttamente e la percentuale di equazioni matematiche risolte correttamente.

Appendice E

Regressions

Working Memory a breve termine:

```

Linear regression                               Number of obs   =       127
                                                F(7, 126)      =       1.36
                                                Prob > F       =       0.2298
                                                R-squared     =       0.0834
                                                Root MSE     =       .18558
    
```

(Std. Err. adjusted for 127 clusters in Task4id)

Percrecall	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GPA	-.0166909	.0134133	-1.24	0.216	-.0432356	.0098537
Income_of_Parents	.008359	.0078898	1.06	0.291	-.0072546	.0239727
Education_of_Father	.0157218	.0167893	0.94	0.351	-.0175039	.0489474
Education_of_Mother	-.0171227	.0208077	-0.82	0.412	-.0583006	.0240552
Age	-.0089642	.0055071	-1.63	0.106	-.0198626	.0019341
MajorDummy	-.0143298	.0341223	-0.42	0.675	-.0818569	.0531973
Sex2	.0397708	.0327328	1.22	0.227	-.0250065	.1045481
_cons	.9377518	.1616473	5.80	0.000	.6178566	1.257647

CRT 1 Score:

```

Logistic regression                               Number of obs   =       127
                                                Wald chi2(8)   =       11.75
                                                Prob > chi2    =       0.1629
Log pseudolikelihood = -68.844219              Pseudo R2      =       0.1382
    
```

(Std. Err. adjusted for 127 clusters in Task4id)

CRT1score	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
Percrecall	1.941634	1.146889	1.69	0.090	-.3062265	4.189494
GPA	-.3071378	.1719565	-1.79	0.074	-.6441664	.0298908
Income_of_Parents	.0526691	.1223655	0.43	0.667	-.187163	.2925012
Education_of_Father	.3608043	.2827602	1.28	0.202	-.1933954	.915004
Education_of_Mother	.2540151	.2452843	1.04	0.300	-.2267334	.7347636
Age	-.0373214	.057542	-0.65	0.517	-.1501016	.0754587
MajorDummy	1.113505	.5271615	2.11	0.035	.0802877	2.146723
Sex2	-.9426849	.467354	-2.02	0.044	-1.858682	-.026688
_cons	-3.86262	2.283483	-1.69	0.091	-8.338165	.6129252

CRT 2 Score:

```

Logistic regression                               Number of obs   =       127
                                                  Wald chi2(8)    =       20.35
                                                  Prob > chi2     =       0.0091
Log pseudolikelihood = -63.551366                Pseudo R2      =       0.1607

```

(Std. Err. adjusted for 127 clusters in Task4id)

CRT2score	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
Percrecall	3.012516	1.287837	2.34	0.019	.4884009	5.536631
GPA	-.5069401	.2479712	-2.04	0.041	-.9929548	-.0209254
Income_of_Parents	.0419647	.1244671	0.34	0.736	-.2019864	.2859158
Education_of_Father	.4880362	.2573384	1.90	0.058	-.0163377	.9924101
Education_of_Mother	-.4845348	.2670943	-1.81	0.070	-1.00803	.0389604
Age	.2026521	.0690645	2.93	0.003	.0672882	.3380159
MajorDummy	-.4893672	.4823476	-1.01	0.310	-1.434751	.4560166
Sex2	-.9489937	.5197997	-1.83	0.068	-1.967782	.0697951
_cons	-6.056511	2.213101	-2.74	0.006	-10.39411	-1.718913

CRT 3 Score:

```

Logistic regression                               Number of obs   =       127
                                                  Wald chi2(8)    =       16.26
                                                  Prob > chi2     =       0.0387
Log pseudolikelihood = -79.168884                Pseudo R2      =       0.1003

```

(Std. Err. adjusted for 127 clusters in Task4id)

CRT3score	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
Percrecall	2.825031	1.04492	2.70	0.007	.7770267	4.873036
GPA	-.3695929	.1615718	-2.29	0.022	-.6862678	-.0529181
Income_of_Parents	.0520884	.1079965	0.48	0.630	-.1595809	.2637577
Education_of_Father	.0137907	.1859552	0.07	0.941	-.3506749	.3782563
Education_of_Mother	-.2342961	.2122309	-1.10	0.270	-.650261	.1816688
Age	.1038102	.0580645	1.79	0.074	-.0099942	.2176146
MajorDummy	.21273	.4326122	0.49	0.623	-.6351743	1.060634
Sex2	-.7398444	.4310032	-1.72	0.086	-1.584595	.1049063
_cons	-2.63887	1.952847	-1.35	0.177	-6.466379	1.188639

Global CRT Score:

```

Multinomial logistic regression      Number of obs    =      127
                                   Wald chi2(24)     =      45.29
                                   Prob > chi2        =      0.0054
Log pseudolikelihood = -143.33976   Pseudo R2       =      0.1405

```

(Std. Err. adjusted for 127 clusters in Task4id)

CRTscore		Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
0	(base outcome)						
1	Percrecall	3.145433	1.22515	2.57	0.010	.7441841	5.546682
	GPA	-.531783	.2013123	-2.64	0.008	-.9263478	-.1372182
	Income_of_Parents	-.0237092	.1200217	-0.20	0.843	-.2589474	.2115291
	Education_of_Father	-.1072664	.236957	-0.45	0.651	-.5716936	.3571609
	Education_of_Mother	-.0328064	.2847676	-0.12	0.908	-.5909406	.5253278
	Age	-.0102611	.095692	-0.11	0.915	-.1978139	.1772917
	MajorDummy	1.450166	.6041982	2.40	0.016	.2659594	2.634373
	Sex2	-1.243882	.5729821	-2.17	0.030	-2.366907	-.120858
	_cons	-.8450818	2.662554	-0.32	0.751	-6.063591	4.373427
2	Percrecall	3.077365	1.401373	2.20	0.028	.330725	5.824004
	GPA	-.7641465	.2938705	-2.60	0.009	-1.340122	-.1881708
	Income_of_Parents	.0499323	.1716258	0.29	0.771	-.2864481	.3863127
	Education_of_Father	.1961894	.2736036	0.72	0.473	-.3400637	.7324426
	Education_of_Mother	.0154828	.3016827	0.05	0.959	-.5758044	.6067701
	Age	.1697918	.0641708	2.65	0.008	.0440194	.2955642
	MajorDummy	.8141789	.5588647	1.46	0.145	-.2811757	1.909534
	Sex2	-.9208424	.5906714	-1.56	0.119	-2.078537	.2368522
	_cons	-5.813228	2.379025	-2.44	0.015	-10.47603	-1.150424
3	Percrecall	4.815932	2.147638	2.24	0.025	.6066394	9.025224
	GPA	-.5175761	.2691049	-1.92	0.054	-1.045012	.0098597
	Income_of_Parents	.074007	.1689971	0.44	0.661	-.2572212	.4052352
	Education_of_Father	.6231258	.3694912	1.69	0.092	-.1010637	1.347315
	Education_of_Mother	-.4515377	.3482384	-1.30	0.195	-1.134072	.230997
	Age	.0719008	.104903	0.69	0.493	-.1337053	.2775068
	MajorDummy	.2349039	.681683	0.34	0.730	-1.10117	1.570978
	Sex2	-1.771889	.7292402	-2.43	0.015	-3.201173	-.3426045
	_cons	-5.572685	4.054379	-1.37	0.169	-13.51912	2.373753

3 Correct Answers Criteria:

```

Logistic regression                                Number of obs   =      127
                                                    Wald chi2(8)    =      12.74
                                                    Prob > chi2     =      0.1211
Log pseudolikelihood = -46.149955                Pseudo R2      =      0.1096

```

(Std. Err. adjusted for 127 clusters in Task4id)

Dscore3	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
Percrecall	2.904915	1.846124	1.57	0.116	-.7134215	6.523251
GPA	-.1916926	.2212428	-0.87	0.386	-.6253205	.2419352
Income_of_Parents	.0628657	.1346828	0.47	0.641	-.2011078	.3268392
Education_of_Father	.5998741	.3371604	1.78	0.075	-.0609482	1.260696
Education_of_Mother	-.4425873	.3046427	-1.45	0.146	-1.039676	.1545014
Age	.0134907	.0940536	0.14	0.886	-.170851	.1978324
MajorDummy	-.4455705	.5975849	-0.75	0.456	-1.616815	.7256744
Sex2	-1.204402	.6318332	-1.91	0.057	-2.442773	.0339681
_cons	-4.369097	3.611032	-1.21	0.226	-11.44659	2.708397

3 Impulsive Answers Criteria:

```

Logistic regression                                Number of obs   =      127
                                                    Wald chi2(8)    =      20.47
                                                    Prob > chi2     =      0.0087
Log pseudolikelihood = -64.422125                Pseudo R2      =      0.2008

```

(Std. Err. adjusted for 127 clusters in Task4id)

Iscore3	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
Percrecall	-3.736323	1.168993	-3.20	0.001	-6.027507	-1.44514
GPA	.6809686	.2061775	3.30	0.001	.2768682	1.085069
Income_of_Parents	.2386754	.1171035	2.04	0.042	.0091567	.4681941
Education_of_Father	-.2824726	.2171651	-1.30	0.193	-.7081084	.1431632
Education_of_Mother	.2602547	.2380907	1.09	0.274	-.2063945	.7269039
Age	-.1110167	.070159	-1.58	0.114	-.2485258	.0264925
MajorDummy	-.6049042	.4710367	-1.28	0.199	-1.528119	.3183107
Sex2	1.363929	.5145737	2.65	0.008	.3553832	2.372475
_cons	.7863722	2.17997	0.36	0.718	-3.48629	5.059034

3 Correct Answers Criteria without Income and Education of Father and Mother:

Logistic regression Number of obs = 127
 Wald chi2(5) = 6.49
 Prob > chi2 = 0.2613
 Log pseudolikelihood = -48.397467 Pseudo R2 = 0.0662

(Std. Err. adjusted for 127 clusters in Task4id)

Dscore3	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
Percrecall	3.347994	1.886672	1.77	0.076	-.349816	7.045803
GPA	-.1076962	.1999856	-0.54	0.590	-.4996607	.2842683
Age	.0144765	.0856988	0.17	0.866	-.15349	.182443
MajorDummy	-.2172762	.5512649	-0.39	0.693	-1.297736	.8631832
Sex2	-1.06624	.6225357	-1.71	0.087	-2.286388	.1539073
_cons	-3.914725	3.207981	-1.22	0.222	-10.20225	2.372802

3 Impulsive Answers Criteria without Income and Education of Father and Mother:

Logistic regression Number of obs = 127
 Wald chi2(5) = 15.75
 Prob > chi2 = 0.0076
 Log pseudolikelihood = -67.082998 Pseudo R2 = 0.1677

(Std. Err. adjusted for 127 clusters in Task4id)

Iscore3	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
Percrecall	-3.411217	1.128556	-3.02	0.003	-5.623146	-1.199288
GPA	.5942062	.1953856	3.04	0.002	.2112575	.977155
Age	-.1031127	.0637755	-1.62	0.106	-.2281103	.0218849
MajorDummy	-.628984	.4782409	-1.32	0.188	-1.566319	.3083509
Sex2	1.260429	.5238023	2.41	0.016	.233795	2.287062
_cons	2.118696	1.718237	1.23	0.218	-1.248988	5.48638

Bibliografia e Fonti Web

- Altmann, S., Falk, A., & Grunewald, A. (2013). Incentives and Information as Driving Forces of Default Effects.
- Baddeley, A. D. (1986). Working memory. Oxford, England: Oxford University Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47–90). New York: Academic Press.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of verbal learning and verbal behavior*, 14(6), 575-589.
- Bar-Hillel, M. (1991). Commentary on Wolford, Taylor, and Beck: the conjunction fallacy?. *Memory & cognition*, 19(4), 412-414.
- Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological bulletin*, 130(4), 553.
- Bosch-Domènech, A., Brañas-Garza, P., & Espín, A. M. (2014). Can exposure to prenatal sex hormones (2D: 4D) predict cognitive reflection?. *Psychoneuroendocrinology*, 43, 1-10.
- Brañas-Garza, P., Kujal, P., & Lenkei, B. (2015). Cognitive Reflection Test: Whom, how, when.
- Campbell, R. D., & Bagshaw, M. (2008). *Human performance and limitations in aviation*. John Wiley & Sons.
- Chaiken, S., & Trope, Y. (1999). Dual-process theories in social psychology. New York: Guilford Press.
- Chamberlin, E. H. (1948). An experimental imperfect market. *The Journal of Political Economy*, 95-108.
- Conway, A. R. A., & Kane, M. J. (2001). Capacity, control and conflict: An individual differences perspective on attentional capture. In C. L. Folk, B. S. Gibson, & S. Bradley (Eds.), *Attraction, distraction and action: Multiple perspectives on attentional capture*. Advances in psychology (pp. 349–372). New York: Elsevier.
- Conway, A. R. A., Kane, M. J., Bunting, M.F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user’s guide. *Psychonomic Bulletin and Review*, 12, 769–786.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185.
- Cowan, N. (2005). Working memory capacity. Hove, East Sussex, UK: Psychology Press
- Cueva Herrero, C., Iturbe-Ormaetxe Kortajarene, I., Mata-Pérez, E., Ponti, G., Sartarelli, M., & Yu, H. (2015). *Cognitive (Ir) reflection: New Experimental Evidence*. Instituto Valenciano de Investigaciones Económicas, SA (Ivie).
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, 11(1), 19-23.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol.44, pp.145–199). New York: Elsevier.
- Engle, R. W., Oransky, N. (1999). The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In R. Sternberg (Ed.), *The Nature of Cognition* (pp. 514-555). Cambridge, MA: MIT Press.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309-331.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGrawHill.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 25-42.
- Gilbert, D. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York: Guilford Press.
- Gilbert, D. (1991). How mental systems believe. *American Psychologist*, 46(2), 107–119.
- Gilbert, D. (1999). What the mind’s not. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 3–11). New York: Guilford Press.
- Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin*, 118(2), 248.
- Holt, C. A., Porzio, M., & Song, M. Y. (2015). *Price Bubbles, Expectations, and Gender in Asset Markets: An Experiment*. University of Virginia working paper.
- Hoppe, E. I. and Kusterer, D. J. (2011). Behavioral Biases and Cognitive Reflection. *Economics Letters*, 110(2):97 – 100.
- James, W. (1890). 1950. *The principles of psychology*.
- Kahneman, D. (2000b). A psychological point of view: Violations of rational rules as a diagnostic of mental processes (commentary on Stanovich and West). *Behavioral and Brain Sciences*, 23, 681–683.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720.

- Kahneman, D., & Smith, V. (2002). Foundations of Behavioral and Experimental Economics. *Nobel Prize in Economics Documents, 1*.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189–217.
- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working-memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp.21–48). New York: Oxford University Press.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189.
- La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(6), 1118.
- Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis*. Psychology Press.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review, 63*(2), 81.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. Holt, Rinehart & Winston, New York.
- Oechssler, J., Roeder, A., & Schmitz, P. W. (2009). Cognitive Abilities and Behavioral Biases. *Journal of Economic Behavior & Organization, 72*(1):147–152.
- Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*.
- Savage, L. J. (1954). *The Foundations of Statistics*. New York: John Wiley
- Smith, C. A., & Morris, L. W. (1976). Effects of stimulative and sedative music on cognitive and emotional components of anxiety. *Psychological Reports, 38*(3c), 1187-1193.
- Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and brain sciences, 23*(05), 701-717.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and brain sciences, 23*(5), 645-726.
- Stanovich, K.E., & West, R.F. (2002). Individual differences in reasoning: Implications for the rationality debate? In T. Gilovich, D. W. Griffin, D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 421-440). New York: Cambridge University Press.
- Sternberg, R. J. (2000). The theory of successful intelligence. *Gifted education international, 15*(1), 4-21.
- Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 83–114). Hillsdale, NJ: Erlbaum.
- Toplak, M. E., West, R. F., and Stanovich, K. E. (2011). The Cognitive Reflection Test as a Predictor of Performance on Heuristics-and-biases Tasks. *Memory & Cognition, 39*(7):1275– 1289.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127-154.
- Unsworth, N., & Engle, R.W. (2007). On the division of shortterm and working memory: An examination of simple and complex spans and their relation to higher-orderabilities. *Psychological Bulletin, 133*, 1038–1066.
- Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert & S. E. Fiske (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 446–496). New York: McGraw-Hill.
- Wittman, W. W. (1988). Multivariate reliability theory. In *Handbook of multivariate experimental psychology* (pp. 505-560). Springer US.