



UNIVERSITÀ DI PARMA

**DIPARTIMENTO DI MEDICINA E CHIRURGIA
CORSO DI LAUREA MAGISTRALE IN PSICOBIOLOGIA E
NEUROSCIENZE COGNITIVE**

**EFFICACIA DEL TESTING EFFECT NEL
POTENZIARE LA MEMORIA PER SCENE NATURALI:
CONFRONTO TRA UNO STUDIO SPERIMENTALE IN
LABORATORIO E UNO STUDIO ESEGUITO ONLINE**

Relatore:

Chiar.ma Prof.ssa VERA FERRARI

Controrelatore:

Chiar.ma Prof.ssa DOLORES ROLLO

Laureanda:

MARIA LUISA SILEO

ANNO ACCADEMICO 2020-2021

ABSTRACT

La memoria è quel meccanismo che ci permette di conservare e recuperare le informazioni nel tempo e di salvarle per un uso futuro ogni qualvolta ci risulti utile farlo. Molto spesso però il recupero delle informazioni fallisce. Nella nostra quotidianità siamo circondati da stimoli visivi e la vista rappresenta un senso straordinariamente efficiente per veicolare le informazioni che provengono dal mondo esterno. Le immagini, dunque, stimolano in modo sorprendente la nostra memoria e potrebbero costituire uno strumento di apprendimento molto potente. Il test è tra i metodi più diffusi per la valutazione della performance ed è generalmente costituito da una serie di domande (in formati e modalità differenti) alle quali i partecipanti devono cercare di rispondere nel modo più accurato possibile. Quindi, sebbene l'obiettivo principale del test sia quello di verificare le proprie conoscenze consolidate attraverso uno studio preliminare, dalla letteratura è emerso che il test è anche un efficace strumento di apprendimento e consolidamento mnestico. Per testing effect, si intende quel fenomeno attraverso cui, l'atto di recupero di informazioni dalla memoria produce un consolidamento migliore delle stesse, rispetto al ripassare le informazioni attraverso il semplice ristudio (Roediger & Butler, 2011). La tecnica del testing si è già visto essere molto efficace con materiale verbale a differenza del materiale non verbale che è stato indagato molto poco; per cui, l'obiettivo del presente elaborato è indagare, attraverso il paradigma del testing effect, la memoria per scene naturali.

Se la rievocazione tramite un test esplicito rafforza la traccia, allora ci aspettiamo che ci sia un vantaggio per le immagini testate rispetto a quelle non testate; in alternativa, se la traccia ha già raggiunto un suo

massimo consolidamento attraverso una prima esposizione, allora il testing non produrrà effetti. Inoltre, come modalità per attivare il recupero mnestico, nel presente studio si è andati a manipolare sia l'emozionalità del materiale da ricordare e rievocare (immagini neutre o immagini emozionali), sia le caratteristiche percettive del materiale usato nel test (nitide o sfocate).

Il qui presente studio, iniziato a ridosso dell'emergenza sanitaria, intende essere un'estensione di uno studio già svolto nel laboratorio di Psicofisiologia di Parma volto ad indagare la memoria per scene naturali attraverso l'uso del testing effect. Il recupero di informazioni dalla memoria produce un consolidamento migliore delle stesse, rispetto al ripassare le informazioni attraverso il semplice ristudio, questo fenomeno di potenziamento mnestico è definito in letteratura "testing effect" (Roediger & Butler, 2011). La tecnica del testing, molto efficace con materiale verbale ma poco indagato con materiale non verbale, qui viene ampiamente esplorato attraverso l'utilizzo di 490 immagini, all'interno di tre sessioni sperimentali come il tipico paradigma del Testing effect prevede (Carpenter, 2006): una prima fase di codifica, in cui sono state presentate immagini dinamiche, una di test di riconoscimento mnestico intermedio (fase testing), in cui gli stimoli sono stati presentati in blocchi di immagini integre e degradate, e un test di riconoscimento mnestico finale, condotto a distanza di una settimana, simile al precedente (vecchio/nuovo), con solo stimoli integri. Poiché in letteratura si è parlato dell'importante ruolo della valenza emozionale nella rievocazione di memorie, si è andati a manipolare sì le caratteristiche percettive del materiale usato nel test (ben visibili o sfocate) che l'emozionalità del materiale da ricordare (immagini neutre o immagini emozionali). Tale studio svolto in precedenza in laboratorio riportava risultati

comportamentali ed elettrofisiologici significativi circa il vantaggio nell'uso del testing per il consolidamento delle memorie a lungo termine. Lo scopo del qui presente lavoro, partendo dal medesimo quesito, è stato quello di implementare il paradigma sperimentale usato per poterlo condurre online cercando di ridurre la variabilità interindividuale attraverso l'uso di accorgimenti e metodiche percettive, valutandone quindi la replicabilità.

In accordo con i risultati dello studio sperimentale svolto in laboratorio e con le nostre ipotesi, anche il presente studio ha mostrato che a distanza di una settimana, nel test finale, la traccia mnestica sembra ulteriormente rinforzata se gli stimoli sono stati presentati nella fase intermedia di Testing, anche in versione degradata, rispetto agli stimoli che sono stati visti solo una volta in codifica. Rimane comunque evidente il vantaggio mnestico per il materiale sottoposto a testing nella sua versione integra. Inoltre, è emerso un contributo significativo dell'emozionalità. In conclusione, per quanto riguarda l'esplorazione della memoria a lungo termine con i parametri che si sono voluti indagare, nonostante il numero ristretto di soggetti e sicuramente la necessità di continuare con la sperimentazione comportamentale a distanza, si può parlare di un accordo tra i risultati emersi in laboratorio e quelli analizzati tramite l'esperimento online.

INDICE

1 INTRODUZIONE	pag. 1
1.1 La memoria	pag. 2
1.1.1 Tipi di memoria	pag. 5
1.1.2 Aspetti neurali	pag. 9
1.2 Testing Effect	pag. 15
1.2.1 Apprendimento attraverso la ripetizione	pag. 17
1.1.2 L'effetto test vero e proprio	pag. 24
1.3 Il potere del test.	pag. 26
1.3.1 Materiale verbale e apprendimento associato in coppie di termini	pag. 26
1.3.2 Materiale in prosa, saggi, test a risposta breve o scelta multipla	pag. 28
1.3.3 Test con materiale più complesso	pag. 31
1.4 Testing online.	pag. 36
1.4.1 Implementazione online di esperimenti in laboratorio . . .	pag. 36
1.4.2 Pro e Contro della ricerca online	pag. 37
1.4.3 Sicurezza e standard per la protezione	pag. 39
1.4.4 Costruzione di un esperimento online	pag. 41

2 LA RICERCA	pag. 49
2.1 Obiettivo dello studio.	pag. 49
2.2 Materiali e Metodi.	pag. 50
2.2.1 Soggetti.	pag. 50
2.2.2 Stimoli	pag. 51
2.2.3 Paradigma	pag. 52
2.2.4 Procedura	pag. 59
2.2.5 Questionari	pag. 61
2.2.6 Registrazione dei dati comportamentali	pag. 62
2.2.7 Disegno sperimentale	pag. 62
2.2.8 Risultati	pag. 63
2.2.8.1 Risultati comportamentali nel Test di riconoscimento Old-New	pag. 63
2.2.8.2 Risultati comportamentali nel Test intermedio	pag. 67
3 DISCUSSIONE	pag. 70
APPENDICE (A-B-C-D)	pag. 79
A-Questionari	pag. 80
B-Script di MATLAB	pag. 86
C-Virtual Chinrest	pag. 87
D- Totale soggetti	pag. 88
BIBLIOGRAFIA	pag. 89

1. INTRODUZIONE

Il test è tra i metodi più diffusi per la valutazione della performance dei soggetti. Esso può essere costituito da diverse domande in formati e modalità di presentazione differenti e ai partecipanti viene chiesto di rispondere in modo accurato e preciso. Questa concezione del test è fortemente radicata nella realtà scolastica ed accademica contemporanea come anche in ambiti riabilitativi nel caso di disturbi traumatici come per esempio nel caso del PTSD. Per testing effect, si intende quel fenomeno attraverso cui, l'atto di recupero di informazioni dalla memoria produce un consolidamento migliore delle stesse informazioni, rispetto al ripassarle e rileggerle nella loro versione originale. Quindi, sebbene l'obiettivo principale del test sia quello di verificare le proprie conoscenze consolidate attraverso uno studio preliminare, dalla letteratura è emerso che il test sia anche un efficace strumento di apprendimento e consolidamento mnestico. Il suddetto elaborato si interessa in particolare all'indagine del fenomeno del testing effect nell'ambito della memoria per scene naturali attraverso l'implementazione online. Questi due anni appena trascorsi in cui siamo stati investiti dall'emergenza Covid-19 hanno messo a dura prova anche l'ambito della ricerca scientifica e ha portato grande difficoltà nella sperimentazione in presenza e in laboratori di ricerca. Perciò, per quanto già da molti anni si utilizzino piattaforme online per arruolare soggetti per la ricerca psicologica, questi due anni sono sicuramente stati decisivi per una sua maggior implementazione e raffinazione delle tecniche.

1.1 Memoria

La celebre frase “Cogito ergo sum”, “Penso dunque sono” (René Descartes, Discorso sul metodo, 1637) per quanto rimanga tra le asserzioni più citate e famose della filosofia occidentale, sembrerebbe ad oggi più corretto riformularla e tradurla con "Io sono, dunque penso". In primo luogo, perché le attività mentali dipendono da una specifica parte del corpo, il cervello, e quindi le due sfere non vengono più viste separate come per il filosofo francese; secondo motivo è che “Siamo chi siamo per ciò che conosciamo e ciò che conosciamo dipende dalla nostra capacità di ricordare ciò che abbiamo appreso”.¹

Ogni nostro ragionamento, ogni pensiero e ogni azione compiuta, li dobbiamo alla nostra memoria, alla capacità del cervello di ricordare e di immagazzinare le esperienze dentro di noi per poi essere recuperate. La memoria è per noi l'impalcatura che sorregge la nostra storia personale e ci permette di crescere, agire e fare esperienza del mondo nel corso della nostra vita.

Essendo il testing effect un fenomeno che studiano e testiamo all'interno di un macro-argomento più ampio come la memoria, non possiamo far a meno di contestualizzarlo dedicandogli uno spazio a sé. La memoria è una capacità fondamentale per la sopravvivenza dell'uomo e non solo; è quella capacità che permette l'adattamento e la sopravvivenza all'interno di un ambiente e proprio per questo importante ruolo che svolge, si è cercato per lungo tempo di classificarla, denominarla e racchiuderla all'interno di definizioni che potessero spiegare al meglio il fenomeno. Dalla metà del XX secolo ad oggi, quando ancora risultava

¹ (Memory: From Mind to Molecules, Larry R. Squire, E. Kandel, W. H. Freeman and Company, 1998)

difficile pensare al processo di memoria come indipendente dalla percezione, dal linguaggio e dal movimento, si sono fatti molti passi avanti fino ad arrivare alla classificazione di diversi tipi di memoria, diversi tra loro e per ognuno dei quali, ci sono regioni cerebrali più interessate rispetto ad altre. La definizione propria della memoria che si ritrova sullo Zingarelli cita: «Funzione generale della mente, consistente nel far rinascere l'esperienza passata, attraverso le 4 fasi di: memorizzazione, ritenzione, riconoscimento, richiamo».

Si ritiene che la memoria sia suddivisa in 3 magazzini, ognuno dei quali con. Delle proprie caratteristiche e una propria capacità di ritenere una diversa quantità di informazioni per un periodo variabile di tempo. Secondo il modello classico dei cognitivisti Atkinson e Shiffrin, le seguenti strutture sono:

- *registro sensoriale* (RS): un magazzino che riconosce l'informazione e la mantiene per pochi secondi e che a sua volta si distingue in **registro sensoriale visivo (o memoria iconica)**, che conserva le immagini; **registro sensoriale uditivo (o memoria ecoica)** che conserva i suoni; **registro sensoriale tattile**, che conserva il ricordo di forme percepite attraverso il tatto.

La conservazione dello stimolo è di tipo *percettivo*, e questo fa sì che se ne mantengano le caratteristiche sensoriali. In questa fase intervengono *l'attenzione selettiva* (che fa sì che ci si focalizzi su alcune informazioni e se ne lascino da parte altre meno importanti) e la memoria a lungo termine, la quale, permette il riconoscimento vero e proprio dello stimolo. Se questa prima fase del processo avviene correttamente, l'unità mnestica passa in un'altra struttura:

— *memoria a breve termine* (MBT): magazzino che contiene un numero limitato di informazioni (7 +/- 2 item) per un periodo di tempo che non va oltre poche ore. Detta anche “memoria operativa”, questa mantiene le rappresentazioni correnti, anche se transitorie, di cognizioni rilevanti per un particolare scopo. Consiste di almeno due sottosistemi, uno per le informazioni verbali e uno per quelle visuo-spaziali, il cui funzionamento viene coordinato da un terzo sistema detto **controllo esecutivo**. L’informazione qui viene segmentata ed analizzata e gli elementi isolati vengono conservati per un certo periodo nella MBT.

A questo punto intervengono alcuni processi fondamentali che permettono il passaggio dell’informazione dalla memoria a breve termine a quella a lungo termine: Si tratta del processo detto di *rehearsal*, ovvero reiterazione e se questo avviene, il ricordo può passare in un altro magazzino di memoria:

— *memoria a lungo termine*: struttura dalle capacità praticamente illimitate, sia per quanto riguarda la quantità di informazioni che per quanto concerne il periodo di tempo nel quale l’informazione può essere mantenuta.

Il materiale contenuto nella MLT ha caratteristiche di ordine prevalentemente semantico. Esso è sottoposto a modificazioni, classificazioni secondo un’organizzazione puramente soggettiva. E proprio per questo, anche il modo in cui l’informazione viene memorizzata, trattenuta, riconosciuta ed infine richiamata, risulta essere diverso a seconda della modalità e del formato dell’informazione.

1.1.1 Tipi di memoria

Si ritiene opportuno iniziare con il descrivere in maniera più accurata la prima grande divisione della Memoria. Essa viene divisa in:

Memoria a Breve termine (**MBT**) e Memoria a Lungo termine (**MLT**), in base all'andamento temporale della conservazione e alla capacità del magazzino.

La MBT, anche detta operativa, è la memoria che contiene i nostri pensieri coscienti e le nostre percezioni momento per momento; essa è volatile, ha una durata e una capacità limitata, per questo non facciamo affidamento sulla MBT nello svolgere attività quotidiane quanto piuttosto, ci permette di manipolare rappresentazioni transitorie di informazioni critiche per gli scopi immediati (Russell Revlin, *Psicologia cognitiva, Teoria e pratica*, Trad. Tartari, G. Galfano, Zanichelli, 2014).

La capacità della MBT aumenta solitamente con l'età degli individui fino a raggiungere il massimo grado nella giovinezza, per poi declinare nella vecchiaia. La capacità di questo magazzino, lo *span* di memoria di un adulto, è di circa 7 ± 2 elementi non in relazione tra loro (Miller, 1956). È però possibile mettere in atto delle strategie, come il chunking (il raggruppamento) per aggirare il problema della capacità limitata. Esso consiste nel mettere in relazione gli item, come fossero un unico raggruppamento di info, come fossero una singola unità. Inoltre, ciò che abbiamo appreso e immagazzinato nella MLT può giocare un ruolo cruciale nel favorire la ritenzione delle informazioni nella MBT. Questo perché, affinché parole e immagini possano costituire un raggruppamento, c'è bisogno che queste risultino familiari al soggetto e disponibili nella memoria a lungo termine: se si mettono in relazione queste informazioni con ciò che già si ha in MLT, risulterà più facile rievocare le informazioni

senza dimenticanze. I due magazzini di memoria, quindi, sono in parte sovrapposti e per questo possiamo dire che lo span di memoria è influenzato dalle conoscenze preesistenti: la MLT.

Per calcolare la durata della permanenza delle informazioni nella MBT è stato utilizzato il compito di Brown-Peterson che ha fatto emergere due risultati (Brown, 1958; Peterson e Peterson, 1959). Primo, il numero degli item che può essere conservato nella MBT si riduce rapidamente con il passare del tempo come se, gli item andassero incontro a decadimento (tra i 15 e i 18 sec senza ripetizione). Il secondo risultato mette in evidenza che, la durata degli item nella MBT è in relazione al numero di raggruppamenti che vi sono presenti. Fino a che siamo in grado di prestare attenzione alle informazioni nella MBT, esse possono rimanere là indefinitamente. Tuttavia, senza attenzione, le info vengono perse in circa 18 secondi. Maggiore è il numero di item e raggruppamenti presenti, maggiore sarà la probabilità che essi vengano confusi tra loro. Questa confusione prende il nome di *interferenza* e contribuisce in maniera significativa alla nostra incapacità di rievocare correttamente gli item dalla MBT.

Un tipo di MBT che necessita di menzione è inoltre la **working memory** (Miller, Galanter, Pribram, 1960). Abbiamo visto che la MBT si comporta in modo diverso in funzione del fatto che le diverse informazioni da ricordare vengano presentate visivamente o uditivamente, rapidamente o lentamente, che gli item attivino informazioni già immagazzinate nella memoria a lungo termine, o dalla specifica posizione occupata dagli item in una sequenza di informazioni più ampia. Gli psicologi cognitivi hanno sviluppato una teoria che riguarda i meccanismi sottostanti la MBT per spiegare meglio come essa, ci aiuti ad interagire con il mondo e a raggiungere i nostri obiettivi. Miller e colleghi attribuirono a questa

“memoria di lavoro” una funzione di supporto per il compiere un lavoro cognitivo come il ragionare, l’ascoltare e il prendere decisioni.

Il modello della memoria di lavoro consiste in quattro componenti che interagiscono tra di loro: il loop fonologico, il taccuino visuo-spaziale, il buffer episodico e l’esecutivo centrale (sist. WM, Baddeley e Hitch, 1974/1977). Le prime tre componenti sono strutture con una limitata capacità di memoria che consentono differenti rappresentazioni di ciò di cui facciamo esperienza: suoni, immagini e configurazioni di sequenze che coordinano queste rappresentazioni. L’esecutivo centrale invece, è responsabile del controllo e della coordinazione di questi tre sottoinsiemi, si assicura che ogni evento di cui facciamo esperienza sia codificato e conservato in memoria per essere poi utilizzato.

Molte delle funzioni della WM coinvolgono il lobo parietale e la corteccia prefrontale ma si è anche visto che, ogni sottoinsieme coinvolge aree distinte del cervello ad eccezione del buffer episodico che sembra rappresentato in molte regioni cerebrali poiché esso è coinvolto in molte funzioni della memoria.

La MLT è quella porzione del nostro sistema mnesico che risulta composta da ogni forma di esperienza e conoscenza raccolta nel corso della nostra vita. È costruttiva e ricostruttiva e si basa sulla formazione di una traccia durevole, ha durata e capacità illimitata. Essa comprende un’ampia gamma di informazioni e i ricercatori l’hanno divisa in due sistemi: **la memoria esplicita** (o dichiarativa) e **la memoria implicita** (o procedurale).

La memoria esplicita è quella che comprende ricordi che cerchiamo di custodire e poi di rievocare in modo deliberato e cosciente. Riguarda la conoscenza di fatti relativi a persone, luoghi, cose e al loro significato, sottende all’apprendimento di fatti ed esperienze ed è una forma di

conoscenza flessibile, che può essere richiamata coscientemente ed essere riferita verbalmente. Tale memoria, che coinvolge l'integrazione di input sensoriali multipli, ha bisogno della complessità del sistema nervoso dei mammiferi, per quanto ci riguarda, nel cervello umano.

È duttile: costruita e ricostruita; richiede la capacità di associare insieme molte e diverse informazioni e si divide a sua volta in memoria semantica ed episodica. La prima, localizzata nel neocortex (lobi temporali laterale e ventrale), è quella che comprende fatti e nozioni del mondo; quella episodica custodisce e collega tempi, luoghi ed eventi specifici della vita di un individuo, è una sorta di memoria autobiografica. Le nozioni (autobiografiche) relative ai tempi e ai luoghi richiedono in particolar modo le aree associative dei lobi frontali (corteccia prefrontale): lesioni a queste aree, infatti, causano la tendenza a dimenticare il modo in cui una nozione è stata acquisita e parliamo di *amnesia delle origini o della fonte*. Può essere retrospettiva nel momento in cui recupero un fatto per riviverne l'esperienza, o prospettica, quando devo ricordarmi di fare qualcosa nel futuro. Per entrambe le memorie comunque, semantica ed episodica, gioca un ruolo fondamentale il lobo temporale mediale.

La memoria implicita è quella che comprende ricordi che vengono richiamati in modo inconsapevole, ci permette di essere influenzati dalle esperienze che non possiamo rievocare in modo deliberato. Essa richiede l'attivazione delle aree del cervello che governano sistemi motori e sensoriali specifici per un dato avvenimento appreso. La forma più semplice di memoria implicita può essere studiata negli invertebrati attraverso meccanismi di riflesso di tipo difensivo (studio sull'*Aplysia californica* negli esperimenti di Kandel). Essa si divide in due sottosistemi: la memoria procedurale e quella percettiva. La prima comprende quelle conoscenze acquisite con la pratica che si traducono

poi, sotto forma di prestazioni (come l'andare in bicicletta). Quella percettiva è costituita da schemi basati sulla percezione. Vengono acquisite mediante l'esecuzione di compiti di tipo riflesso e attraverso diverse forme di apprendimento (percettivo, motorio ecc.). La memoria implicita è localizzata nei circuiti percettivi motori e limbici (Principi di Neuroscienze, Eric R. Kandel, J. H. Schwartz , T.M. Jessell, Zanichelli, 2014, V ediz.)

1.1.2 Aspetti neurali

I processi di memorizzazione sono dovuti a modificazioni della capacità di trasmissione sinaptica tra neuroni. Le nuove vie di trasmissione del segnale, generate da queste modificazioni sinaptiche, portano alla costruzione nel sistema nervoso centrale (SNC) di tracce di memoria o tracce mnesiche che, una volta stabilite, possono essere attivate in modo selettivo per rievocare i ricordi.

Perché ci possa essere un ricordo, è importante ci sia una qualche forma di apprendimento; l'informazione cioè deve essere acquisita. Una volta che l'informazione è stata acquisita, essa deve essere mantenuta e conservata in memoria fino a che non ci serve. Infine, questa informazione viene usata, cioè, noi ricordiamo. Questo avviene attraverso il "ripescaggio" dell'informazione dalla memoria per riportarla ad uno stato attivo. Gli studiosi della memoria denominano queste fasi del ricordo con i nomi di *codifica*, *consolidamento*, *conservazione* e *recupero*. Ovviamente, questo processo interessa la memoria esplicita, ovvero quel magazzino che comprende quelle informazioni che possono essere descritte e recuperate consapevolmente. Le quattro fasi più nello specifico sono: la *codifica* che si riferisce al modo in cui un'informazione nuova

viene inserita in un contesto di informazioni precedenti, abbiamo qui un'elaborazione e associazione logica con quanto è già in memoria. La forza della traccia dipende dalla profondità della codifica. Il *consolidamento* è la fase in cui, attraverso l'espressione di geni e la sintesi di proteine, vediamo le variazioni strutturali delle sinapsi. Questo è il meccanismo che rende più persistenti le informazioni conservate. La *conservazione* fa riferimento a quei siti neurali nei quali la memoria viene conservata nel tempo. Il *recupero*, infine, è la fase di ricostruzione. Perché esso avvenga è necessario sia presente un appropriato suggerimento che in qualche modo "riattivi" gli elementi focali dell'informazione da ricordare. Non sono le caratteristiche della traccia mnemonica in quanto tali a determinare il ricordo, quanto piuttosto la compatibilità tra le proprietà della traccia e le caratteristiche dell'informazione fornita al recupero. Questa interazione tra informazione immagazzinata e informazione presente al recupero è al centro di un principio che prende il nome di *principio della specificità della codifica* (Solomon Et al., Biologia, VI edizione, Editore Napoli).

Per la memoria esplicita è necessaria la presenza di lobi temporali della corteccia cerebrale, incluso l'**ippocampo**, **area implicata nell'apprendimento esplicito, fornisce informazioni contestuali, spaziali e temporali relative alla traccia mnemonica (es. situazioni di pericolo)**. L'ippocampo, nei vertebrati superiori, è la componente del sistema temporale mediale deputata al consolidamento e alla conservazione delle tracce di memoria dichiarativa. Gli esseri umani possiedono due ippocampi, uno per ogni emisfero del cervello, ha una forma curva e convoluta. Esso fa parte della formazione dell'ippocampo, parte del telencefalo, situata nella regione ventromediale di ciascun lobo temporale, è costituita dall'ippocampo propriamente detto, dalla

corteccia entorinale, dal subicolo e dal giro dentato; svolge un ruolo importante nella memoria a lungo termine e nella navigazione spaziale.

Nell'uomo, il coinvolgimento di queste aree è stato osservato in relazione alle conseguenze della distruzione chirurgica dell'ippocampo nel tentativo di trattare le crisi epilettiche (il caso più famoso è quello del paziente H.M. che ha subito l'ablazione bilaterale dell'ippocampo e dei 2/3 anteriori di tutto il lobo temporale mediale). Il paziente, in seguito all'operazione presentava una grave amnesia e non ricordava cosa fosse accaduto sia prima che dopo l'intervento, anche in un lasso temporale di anni. L'ippocampo è costituito da un circuito tri-sinaptico che agisce attraverso la via diretta (o perforante) che porta informazioni dalla corteccia entorinale ai neuroni granulari del giro dentato; la via delle fibre muscoidi composta dagli assoni dei neuroni granulari che innerva i neuroni piramidali della regione CA3, e la via delle collaterali di Schaffer, formata da collaterali assoniche dei neuroni piramidali della regione CA3, che innervano i neuroni piramidali della regione CA1 (Principi di Neuroscienze, Eric R. Kandel, J. H. Schwartz, T.M. Jessell, Zanichelli, 2014, V ediz.).

Nel 1973 Bliss e Lomo scoprirono che uno stimolo ad alta frequenza in una qualsiasi di queste tre vie dava origine ad un consistente rinforzo delle sinapsi stimulate: questa facilitazione è stata definita potenziamento a lungo termine (LTP). Ha durata di qualche ora ma può essere prolungata fino a giorni attraverso scariche ripetute (LTP tardivo). Il long term potentiation è una forma di plasticità sinaptica che consiste in un aumento a lungo termine della trasmissione del segnale tra due neuroni, in seguito a stimolazione tetanica del neurone presinaptico (con aumento dell'ampiezza del EPSP nei neuroni bersaglio). I cambiamenti dell'attività sinaptica nell'ippocampo e l'LTP sono stati

identificati come i meccanismi cellulari più probabili tra quelli che sottendono l'acquisizione della memoria. Se la plasticità sinaptica può essere definita come la capacità di cambiamento della forza delle sinapsi, l'LTP è un persistente incremento della forza sinaptica indotta da un breve stimolo elettrico ad alta frequenza delle fibre afferenti o l'attivazione simultanea di neuroni pre e post-sinaptici. Il fenomeno è dovuto ad un protratto ingresso di Ca^{++} attraverso canali NMDA, i quali si aprono solamente in presenza di persistente. (R. Carroll, Elsevier Masson, fisiologia, a cura di Luciano Zocchi, 2012). Il processo inizia con la trasmissione sinaptica che attiva i recettori metabotropici del glutammato e i recettori AMPA, poi scariche ad alta frequenza inducono l'LTP attivando anche i recettori NMDA grazie all'eliminazione del blocco del magnesio (Mg $^{+}$). L'ingresso di ioni calcio (Ca_{2+}) produce reazioni di trasduzione del segnale che fanno aumentare la sensibilità dei recettori AMPA, migliorando l'efficienza di risposta della cellula post-sinaptica al rilascio di neurotrasmettitore dal presinaptico.

Altri effetti a lungo termine di questo processo sono il rilascio da parte di neuroni post sinaptici di un fattore neurotrofico che agisce in maniera retrograda in modo da aumentare l'efficienza neurotrasmettitoriale del neurone presinaptico; l'aumento del numero e delle dimensioni dei contatti sinaptici (con aumento e crescita delle spine dendritiche), quindi vera e propria modificazione della morfologia della struttura sinaptica; grazie a meccanismi di questo tipo, la complessità del nostro sistema nervoso aumenta con l'esperienza.

Le tre caratteristiche fondamentali dell'LTP, che si devono realizzare affinché avvenga la conservazione della memoria sono: la cooperatività, l'associatività e la specificità (o sinapsi-specifica).

La **cooperatività** si ottiene quando si stimolano un gran numero di fibre contemporaneamente. Nel post-sinaptico si somma così il contributo depolarizzante di ogni assone.

L'**associatività** rappresenta la condizione secondo la quale devono entrare in attività *simultaneamente* sia il neurone presinaptico che post-sinaptico. La **specificità** sta a significare il fatto per cui solo l'assone stimolato tetanicamente viene indotto verso LTP.

L'LTP è sicuramente il sistema più utilizzato per registrare informazioni per giorni o settimane all'interno di questo "magazzino temporaneo" che è l'ippocampo, ma, essendo un magazzino anche "plastico" deve avere la possibilità di registrare temporaneamente queste info per poi cancellarle. Occorre così un fenomeno inverso in grado di cancellare l'LTP: questo fenomeno è noto come depressione a lungo termine (LTD). Il meccanismo è quello di stimolare il neurone presinaptico con bassa frequenza per tempi prolungati per riuscire ad indurre LTD nel neurone post-sinaptico. Tuttavia, per sinapsi precedentemente potenziate attraverso LTP, sono necessarie diverse stimolazioni per ritornare alle condizioni di base iniziali (Principi di Neuroscienze, Eric R. Kandel, J. H. Schwartz , T.M. Jessell, Zanichelli, 2014, quinta ediz.).

Ogni stimolazione a bassa frequenza di LTD si abbassa un po' la soglia di EPSP raggiunta con la LTP. Questi due meccanismi, comunque, sembra possano susseguirsi un numero infinito di volte, cosa fondamentale per la nostra plasticità mnemonica.

La memoria implicita o non dichiarativa invece, a differenza di quella esplicita di cui abbiamo trattato fino ad ora, si serve di un apprendimento procedurale che prevede una risposta motoria ad una stimolazione sensoriale. L'area implicata nell'apprendimento e nella

memoria implicita è l'amigdala. Essa, attraverso proiezioni ad altre aree cerebrali, tra cui l'ippocampo, facilita i processi di memoria per stimoli a contenuto emotivo (come la paura).

Questo tipo di apprendimento si suddivide in: apprendimento non associativo e apprendimento associativo. Il primo è chiamato così perché prevede che una sola stimolazione sia sufficiente a generare apprendimento e si divide in: abitazione e sensibilizzazione. L'abituazione è una risposta che si riduce progressivamente al progredire della stimolazione, la sensibilizzazione è un processo sostanzialmente inverso: se si viene sensibilizzati ad uno stimolo, il suo effetto resisterà più a lungo nonostante le successive presentazioni. Nell'apprendimento associativo invece è necessario si venga a creare un legame tra eventi in modo da guidare il comportamento. Gli esempi migliori di questo apprendimento sono il condizionamento classico e quello operante.

Studi sul condizionamento delle risposte di evitamento hanno messo in evidenza che uno dei ruoli dell'amigdala è quello di stabilire una relazione tra stimoli esterni e conseguenze in termini di ricompense e punizioni. Nel condizionamento pavloviano, l'amigdala integra le informazioni relative allo stimolo condizionato (SC) con quelle relative allo stimolo incondizionato (SI); una lesione al nucleo laterale o centrale dell'amigdala impedisce che questo condizionamento avvenga. L'attivazione dell'amigdala è elevata durante l'accoppiamento SN (neutro) ed SI, ma una volta condizionato lo stimolo, l'attività si riduce rapidamente nel tempo (Psicologia generale: la scienza della mente e del pensiero, Passer, Holt, Sutherland, Smith, McGraw-Hill Education, 2015).

Attraverso studi di fMRI si è visto che il condizionamento coinvolge altre aree oltre all'amigdala: tra queste troviamo l'ippocampo, come già detto, la corteccia del cingolo e l'insula.

1.2 Testing Effect

Per *testing effect*, si intende, quel fenomeno per cui l'atto di recuperare le informazioni dalla memoria produce una conservazione migliore rispetto all'andare a ristabilire le stesse informazioni per un tempo uguale, attraverso il semplice studio (Roediger & Butler, 2011). La ricerca sulla memoria ha ripetutamente dimostrato un'interdipendenza tra i processi di codifica di nuove informazioni, immagazzinamento nel tempo e accesso postumo attraverso il recupero. Una chiara dimostrazione di questo è appunto l'effetto test, ovvero la scoperta che recuperare informazioni dalla memoria può, in più di alcune circostanze, rafforzare la memoria di queste informazioni che vengono recuperate (Roediger & Butler, 2011; Roediger & Karpicke, 2006a). E sebbene l'effetto positivo e migliorativo del test di memoria sulla ritenzione sia noto già da tempo (come si è visto nelle prime indagini ed esperimenti empirici e non di Abbott, 1909; Gates, 1917; Spitzer, 1939), la ricerca su tale argomento continua a crescere esponenzialmente nel corso del decennio corrente (vedi Rawson & Dunlosky, 2011).

L'effetto del test si presenta come un fenomeno solido e dimostrato attraverso una grande varietà di materiali, quali elenchi di parole singole (ad es. Carpenter & DeLosh, 2006; Rowland & DeLosh, 2014b; Rowland, Littrell-Baez, Sensenig e DeLosh, 2014 ; Zaromb & Roediger, 2010), liste di parole accoppiate (ad esempio, Allen, Mahler, & Estes, 1969; Carpenter, 2009; Carpenter, Pashler, & Vul, 2006; Carrier & Pashler, 1992; Pyc & Rawson, 2010; Toppino & Cohen, 2009), brani in prosa (es. Glover, 1989; Roediger & Karpicke, 2006b) e materiali non verbali (es. Carpenter & Pashler, 2007; Kang, 2010). E nonostante questa grande variabilità di materiali utilizzati appartenenti a domini differenti

(semantici e non), gli studi sull'effetto test possono essere riconosciuti come aventi poche e determinanti fasi discrete esposte di seguito:

Fase 1) Le informazioni, di qualsiasi tipo esse siano (il materiale oggetto di studio) vengono presentate ai partecipanti per una fase di apprendimento iniziale.

Fase 2) Dopo l'apprendimento iniziale, si verifica una fase intermedia durante la quale le informazioni possono:

- essere ripresentate per un ulteriore studio (condizione di restudy);
- sottoposte a un test di memoria (condizione di test; con o senza feedback correttivo);
- non ripresentate o testate (nessun test o condizione di solo studio);

Fase 3) Dopo un intervallo di ritenzione, viene fornita una valutazione finale della memoria per le informazioni che sono state precedentemente apprese.

L'effetto del test si riferisce alla presa di coscienza che le informazioni sottoposte a un test nella fase intermedia di un esperimento, vengono ricordate meglio in una fase di valutazione finale rispetto alle informazioni riprese nella condizione di restudy, o a quelle non ripresentate affatto dopo l'apprendimento iniziale. Prima ancora che venissero raccolte delle prove sperimentali sul testing effect, Francis Bacon nel 1620 parlò dell'utilità del test come buona strategia di apprendimento. "Se leggi un pezzo di testo venti volte, non lo imparerai a

memoria così facilmente come se lo leggessi dieci volte cercando di recitarlo di tanto in tanto e consultando il testo quando la tua memoria viene meno" [Bacon 2000 (1620)] Successivamente verso la fine del XIX secolo anche William James descrive questo effetto nel suo libro "I principi della Psicologia" in cui dichiara come la ripetizione attiva rispetto a quella passiva porti ad un vantaggio mnestico (James 1890); e anche CA Mace più tardi (Psicologia dello studio,1932) affermerà: <<*quando si alternano atti di lettura e atti di richiamo, cioè quando ogni lettura è seguita da un tentativo del soggetto per ricordare gli elementi, l'efficienza dell'apprendimento e della conservazione è enormemente migliorata*>>.

1.2.1 Apprendimento attraverso la ripetizione

Prima di parlare di effetto test vero e proprio, partiamo col parlare di due concetti importanti che si incontrano quando si tratta il tema della memoria: quello di chunking e di ripetizione come mezzi attraverso i quali le persone possono ricordare più agevolmente informazioni, fatti ed eventi.

Il primo stadio che caratterizza il funzionamento della memoria è la codifica: una semplice info entra nel nostro sistema e assieme ad una rete fitta di altre informazioni viene aggiunta e registrata; è la fase che si attiva fin dai primi attimi in cui si presenta il materiale da elaborare ed apprendere. Essa è una fase molto importante in quanto, se la soglia d'attenzione che il soggetto presta allo stimolo non è adeguata o il suo stato psicofisico ed emotivo è alterato, il materiale potrebbe non essere codificato con successo e l'informazione non verrà trasformata in una rappresentazione mentale idonea all'immagazzinamento.

Quando invece tutte queste condizioni sopra citate si verificano, la seconda fase, cioè l'immagazzinamento, ha luogo. Essa è anche chiamata fase di ritenzione ed è caratterizzata dal consolidamento della traccia mnemonica, codificata in entrata, permettendone così la conservazione in un magazzino. Tale processo è supportato e garantito dalla RIPETIZIONE o *reiterazione del materiale*, la quale fa sì che il materiale non venga perso.²

*Nel lavoro di Henderson e James "Consolidating new words from repetitive versus multiple stories: Prior knowledge matters", il cui titolo tradotto è **Consolidare nuove parole con ripetizione di storie vs molteplici storie: le conoscenze precedenti sono importanti**, i due ricercatori del Dipartimento di Psicologia di York hanno cercato di confermare l'idea secondo la quale sono le **conoscenze precedenti** a far gran parte del lavoro di immagazzinamento delle informazioni durante la memorizzazione delle informazioni che arrivano al nostro sistema.*

Lo studio ha osservato che i bambini con conoscenze vocabolarie superiori mostravano un rafforzamento consolidato durante la notte, soprattutto quando, in storie differenti, si incontrano nuove parole.

I bambini di età compresa tra i 10 ei 11 anni (N = 42) sono stati esposti a due set di otto parole (ad esempio, "coccodrillo"), con un set di otto parole incorporato nella stessa storia presentata due volte (ripetizione) e l'altro set di parole presentato invece in due storie differenti.

I bambini con conoscenze vocabolarie superiori hanno mostrato maggiori guadagni durante la notte sia in conoscenza fonologica che semantica esplicita quando nuove parole avevano avuto luogo in più storie, diverse.

² (Alan Baddeley, La memoria come funziona e come usarla, Laterza, Roma-Bari, 1993)

Tuttavia, quando nuove parole erano state incontrate in storie ripetitive (la stessa storia ripetuta più volte), le conoscenze esistenti non esercitavano alcuna influenza sul consolidamento di conoscenze fonologiche esplicite e avevano un impatto negativo sul consolidamento della conoscenza semantica.

Ventiquattro ore dopo l'esposizione della storia è stata osservata un'integrazione lessicale piuttosto debole delle nuove parole imparate in storie multiple, rispetto a quelle acquisite attraverso storie ripetute, potendo affermare quindi che: sebbene il *consolidamento di nuove conoscenze* vocabolari esplicite acquisite attraverso contesti multipli sia supportato da conoscenze precedenti, *l'integrazione lessicale* potrebbe trarre maggiore vantaggio dalla ripetizione.

Lo storytelling orale durante l'infanzia è importante per lo sviluppo sociemozionale e conferisce significativi vantaggi educativi soprattutto per quanto riguarda l'alfabetizzazione. Le esperienze di alfabetizzazione domestica e anche l'esposizione a libri di fiabe sono per i bambini mezzi fondamentali per lo sviluppo del loro vocabolario. Ciononostante, la capacità dei bambini di raccogliere nuove parole dall'ascolto di storie è molto variabile e cambia anche in base alle loro conoscenze pregresse. Attraverso studi sull'apprendimento di parole pronunciate è stato visto il ruolo attivo del sonno. Questo consolidamento *offline* risulta importante per gli aspetti più fonologici piuttosto che semantici dell'apprendimento delle parole. Henderson ha riscontrato miglioramenti durante la notte nella capacità dei bambini di ricordare le definizioni delle parole precedentemente sconosciute ma, non sono emersi risultati positivi circa l'esprimere un giudizio o dare significato alle nuove parole appena apprese.

Lo studio riportato si propone di scoprire le condizioni ottimali per l'apprendimento di nuove parole dalle storie. Bambini di 10 e 11 anni sono stati esposti a due serie di otto parole nuove ascoltando storie fittizie. Un set di queste parole è stato presentato nella stessa storia per due volte (condizione di contesto ripetitivo) e l'altro set è stato presentato in due storie diverse (condizione di contesto multiplo), vedendo anche se il consolidamento notturno, portasse benefici o meno. I bambini sono stati esposti tutti a condizioni di storie multiple e ripetitive in una singola sessione e si è poi misurato il loro apprendimento poco dopo l'esposizione e a ventiquattro ore di distanza per far sì che l'influenza contestuale fosse minima. I bambini hanno imparato nuovi termini, simili a parole inglesi familiari (es. *dolpheg-dolphin*) per misurarne la capacità di recuperare la conoscenza semantica esplicita ma anche la capacità di integrazione lessicale, entrambe all'interno dello stesso compito (modello della coorte di Marslen-Wilson (1978), Schacter D.L., Gilbert D.T., Wegner D.M.)³: una volta che "dolpheg" è stato integrato nel lessico, sentendo la parola "dolphin", i bambini dovrebbero co-attivare entrambi i termini (Effetto di concorrenza lessicale). Se i nuovi item iniziano a competere con le parole d'ordine (quelle familiari dall'inizio) per il riconoscimento, allora i giudizi e i significati attribuiti a questi dovrebbero aumentare rispetto a quelli dati alle parole familiari.

I ricercatori hanno ipotizzato anche un potenziamento dell'apprendimento e/o consolidamento e integrazione per le parole incontrate in storie multiple (non ripetute) per uno stesso gruppo d'età; e che qualsiasi vantaggio dell'apprendimento fosse positivamente associato alla conoscenza già esistente del vocabolario.

³ Psicologia generale. Zanichelli, Bologna, 2010

Il disegno di ricerca e i metodi utilizzati in questo studio, hanno coinvolto 42 bambini di età compresa tra i 10 e gli 11 anni, appartenenti allo stesso contesto socioeconomico. L'età è stata scelta per poter confrontare lo studio con altri precedenti (Wilkinson e Houston-Price, 2013) e perché il vocabolario a questa età è già abbastanza ampio. Il campione scelto è stato sottoposto a test standardizzati per misurare l'abilità non verbale, il vocabolario espressivo, comprensione e produzione di definizioni di parole a vari livelli di difficoltà: il campione si è confermato rappresentativo. Gli stimoli usati nello studio sono parole, parole nuove, storie e illustrazioni, tutti forniti come materiale online.

Le parole nuove sono state presentate in liste di coppie di stimoli; la coppia comprendeva una parola nuova parlata ("crocodol") e una parola base familiare ("coccodrillo"). Per comprendere anche la misurazione della competizione lessicale nell'attività di animazione, sono state prese inoltre due ulteriori liste di otto punti, avendo così due liste sperimentali di parole base e due liste di parole di controllo. La metà delle parole in ciascuna lista erano oggetti non viventi e l'altra metà erano animali, ogni nuova parola era abbinata ad un oggetto/animale (Weighall, 2017) ed ognuna aveva tra le tre e le cinque caratteristiche semantiche. La maggior parte di queste erano compatibili con l'oggetto/animale familiare, altre erano caratteristiche invece nuove.

Le storie sono state presentate audio visivamente tramite iPad. Sono state scritte dagli sperimentatori quattro storie: due a tema "I viaggi di Lucy in uno zoo alieno" e contenevano le parole della lista 1; e le altre due a tema "Le avventure di Jack sulla luna", contenenti le parole della lista 2. Ogni nuova parola è stata presentata tre volte durante il corso di tutta la storia e non sono stati forniti espliciti significati di queste parole. Ogni storia inoltre è stata accompagnata da quattro illustrazioni a colori, essa

rimaneva sullo schermo per tutta la durata del segmento della storia a cui era stata assegnata.

Dopo aver ascoltato le due ripetizioni della storia nella condizione dei contesti ripetitivi e dopo aver ascoltato ciascuna storia nella condizione di più contesti, i bambini hanno risposto a tre domande per verificare la loro comprensione. Ogni domanda è stata valutata come 1 o 0 (massimo di 3 per la condizione di più contesti). I bambini hanno raggiunto i punteggi massimi per entrambe le condizioni, suggerendo così l'idea che stessero comprendendo con successo le storie (media contesti=2.79, SD=1.32).

La sessione di test si è rivelata però essere un predittore negativo dei tempi di reazione complessivi ($p=0.03$), evidenziando giudizi rapidi dovuti alla sola pratica del compito. Nessun altro predittore del modello offre un forte supporto agli effetti di competizione lessicale. Contrariamente alle ipotesi iniziali, è emerso solo un rallentamento *marginale* delle risposte alle parole base (familiari) rispetto a quelle di controllo.

È stato fatto fare un compito di recupero semantico e integrazione lessicale esponendo i bambini all'ascolto di parole, parole base familiari, parole base di controllo ed è stato chiesto loro di decidere con un sì/no se si trattasse di un animale o meno (16 parole nuove, 16 parole familiari, 16 di controllo più altre 16 parole di riempimento per bilanciare il numero di parole incontrate o meno nelle storie). I bambini sono stati sottoposti poi ad un compito di richiamo guidato per misurare la capacità di recuperare le forme fonologiche delle parole ottenendo 1 per la corretta produzione di ogni nuova parola e 0 per ogni risposta errata. Hanno inoltre sostenuto un altro compito per misurare la loro capacità di esprimere il significato delle parole (ogni parola aveva un punteggio massimo di 3).

Ciò che è emerso in definitiva è che: per quanto riguarda il richiamo fonologico, la sessione di test si è rivelata un'importante fattore predittivo delle prestazioni di richiamo a distanza, l'accuratezza era migliore il giorno seguente rispetto al giorno dell'apprendimento e anche la conoscenza del vocabolario influiva sulle prestazioni generali. L'interazione tra vocabolario e sessione di test però, risulta significativa solo nella condizione di più contesti. Solo in questa condizione cioè, la capacità del vocabolario risultava predittore positivo per le prestazioni nella seconda sessione di test.

Nel presente studio non vi sono state prove evidenti del fatto che le parole del romanzo fossero state integrate con le conoscenze lessicali esistenti in seguito all'esposizione alle storie; tuttavia, ci sono alcune *prove deboli che suggeriscono che l'integrazione lessicale potrebbe essere più probabile a seguito dell'esposizione a storie ripetitive*. Vale a dire, 24h dopo l'esposizione della storia c'era stato un leggero rallentamento delle risposte alle parole familiari rispetto a quelle nuove, che invece erano state integrate nelle conoscenze lessicali (limitatamente però, come si è detto, alla condizione di contesto ripetuto). Sebbene siano dati deboli che richiedono sicuramente altri studi e condizioni sperimentali, questi risultati potrebbero suggerire che la ripetizione possa esercitare un vantaggio per la nuova integrazione lessicale delle parole. Il recupero fonologico automatizzato può trarre beneficio da un allenamento ripetitivo, come fosse la riattivazione di una traccia di memoria. Al contrario, i compiti che implicano un recupero esplicito di conoscenze semantiche ma anche fonologiche, traggono maggior vantaggio da legami più ricchi con la conoscenza semantica e quindi con il vocabolario. Ad avvalorare questi risultati sulla ripetizione come mezzo utile per l'integrazione lessicale è lo studio di Henderson (2015) in cui ai bambini

veniva chiesto di ripetere le parole del romanzo non rispondendo a delle domande dopo l'ascolto di storie ma, ad alta voce in isolamento: l'integrazione di queste parole nel loro vocabolario è stata più forte che in altre condizioni.

In conclusione, all'esamina di questo studio si può dire che: esposizioni più estese che incorporano istruzioni esplicite affiancate a narrazioni di storie, possono essere necessarie per un solido apprendimento a lungo termine, coerenti con il fatto che, la semplice esposizione dei bambini ai libri non è sufficiente a potenziare lo sviluppo del loro linguaggio orale. In particolare, l'accento è stato posto sull'importanza delle conoscenze pregresse per l'acquisizione di nuovi vocaboli in contesti multipli, forse per via delle tante connessioni che si venivano a formare tra conoscenza nuova ed esistente (Lewis e Durrant, 2011). Tuttavia, sebbene come è stato detto questa conoscenza non è sufficiente e non ha un impatto positivo sul *consolidamento* delle conoscenze esplicite acquisite attraverso storie ripetitive, il divario tra bambini con un lessico più ricco e quelli con un vocabolario più ristretto potrebbe essere ridotto attraverso l'esposizione a storie ripetitive che andrebbe ad incrementare, almeno in parte, l'integrazione lessicale.

La ripetizione risulta così essere uno strumento mentale importantissimo (Henderson, L.M.; James, E. (2018).

1.2.2 L'effetto test vero e proprio

Perciò evidenze sull'importanza di questo fenomeno vennero già esposte prima ancora del XXI secolo, momento nel quale, riconosciamo il massimo interesse per questa pratica. Prime vere evidenze empiriche del fenomeno si hanno però nel 1909 grazie ad un lavoro di Edwina E. Abbott

che, con il suo esperimento “On the Analysis of the Memory Function in Orthography” condotto nel laboratorio psicologico dell'Università dell'Illinois, arrivò alla conclusione che l'opportunità di richiamo, durante o immediatamente dopo il processo di apprendimento, era di grande beneficio per l'individuo. Venne indagato, quando viene dato un tempo definito in cui imparare del materiale, se è di maggior vantaggio passare tutto il tempo nella percezione effettiva dello stesso, o parte del tempo nella percezione e parte del tempo nel richiamo e, inoltre, se il richiamo dovrebbe essere intervallato dalla percezione o dovrebbe seguirla immediatamente dopo o essere presentato dopo un certo lasso di tempo. Questo approccio in cui si alterna la lettura, l'auto-test e il feed back (sotto forma di rilettura quando i tentativi di recupero falliscono) è stato coniato come *recitazione*. Vedendo l'uso di questa pratica nell'istruzione, i test danno l'opportunità di apprendere dal feedback che lo studente stesso riceve sulla performance del test, specialmente quando quel feedback è elaborato e significativo, come nel caso della tecnica di valutazione formativa presso le istituzioni scolastiche. Inoltre, se gli studenti si mettono alla prova periodicamente mentre studiano, possono usare il risultato di questi test per guidare il loro studio futuro verso il materiale che non hanno ancora imparato. Il sostenere un test, funziona come una sorta di promozione dell'apprendimento futuro rendendo più efficace lo studio successivo o la codifica; quindi, il test migliora l'apprendimento attraverso questo processo di mediazione. Il valore relativo di questo tipo di metodo di apprendimento, inoltre, si è visto dipendere anche dal tipo di memoria dell'individuo. Ciò che può risultare vantaggioso per un individuo dalla mente visiva potrebbe non esserlo per uno del tipo uditivo-motorio. Perciò ci sono molti fattori da dover tenere in considerazione nell'analisi del testing effect.

1.3 Il potere del Test

1.3.1 Materiale verbale e apprendimento associato in coppie di termini

Uno dei primi studi che dimostra il potere della recitazione è un esperimento di Gates (1917, *Recitation as a factor in memorizing. Archives of Psychology, 40*, Pp. 104). Nell'esperimento si erano dati a un gruppo di bambini 9 minuti per imparare del materiale nuovo (sillabe senza senso o biografie). Gli sperimentatori hanno manipolato la percentuale di tempo che è stato impiegato nella recitazione (0, 20, 40, 60, o 80% del periodo di apprendimento di 9 minuti, con il tempo rimanente speso nella lettura del materiale). Si è visto che le prestazioni su un test finale (dato dopo il periodo di 9 minuti) variavano in funzione della quantità di recitazione, anche se l'effetto differiva tra i materiali. Per le sillabe senza senso c'era un effetto maggiore, mentre per le biografie gli effetti sul richiamo erano misti. Vieni poi visto un effetto ancora più sorprendente quando ai partecipanti non viene data la possibilità di tornare al materiale originale per studiarlo nuovamente, tra la fase di test iniziale e quella di test finale. Fare uno o più test intermedi (in cui il feedback è solo personale e non vengono fornite delle risposte corrette da parte dello sperimentatore) si è visto che migliora la performance e aiuta, almeno parzialmente, il processo di dimenticanza. Minore è il tempo trascorso tra lo studio e il test, migliore sarà la performance. Fare un test di memoria effettivamente attenua l'oblio della stessa. (Spizer Studies in retention. *Journal of Educational Psychology, 30*(9), 641–656). E questo principio per cui inserire un test di richiamo nella sequenza di apprendimento aumenta la probabilità che l'allievo ricordi qualcosa durante un test successivo è stato dimostrato

empiricamente da molti altri ricercatori, tra cui Izawa (1966, 1970), Donaldson (1971), Madigan e McCabe (1971), Young (1971), Bartlett e Tulving (1974), Modigliani (1976) e Bartlett (1977).

Per esempio, in uno di questi studi, Bartlett e Tulving, in *Effects of temporal and semantic encoding in immediate recall upon subsequent retrieval* (1974, *Journal of Verbal Learning & Verbal Behavior*, 13(3), 297–309) hanno fatto imparare ai soggetti una lista di coppie di parole associate a numeri, andando a studiare il fenomeno in una condizione di domini differenti. Hanno condotto, infatti, due esperimenti con 72 studenti universitari, in ciascun esperimento i soggetti ascoltavano elenchi di 6 coppie di cifre e parole. Dopo ogni elenco tentavano di ricordare una delle parole in risposta a un segnale temporale o semantico; la ritenzione delle coppie di parole-numeri da parte dei soggetti fu poi testata con un test di richiamo libero o un test di riconoscimento. Le parole inizialmente non testate hanno mostrato un forte effetto negativo di recency nel richiamo e nel riconoscimento ritardati, suggerendo che l'effetto derivi da processi di codifica che terminano con la presentazione di ciascun elenco. Il richiamo e il riconoscimento ritardati delle parole inizialmente testate, sono stati influenzati dalla modalità di recupero iniziale solo per le parole dalle posizioni di input 5 e 6, essendo il cuing semantico iniziale più efficace del cuing temporale.

Prima del test finale, Bartlett e Tulving diedero ai soggetti un test di richiamo su alcune delle coppie parole-numeri. Ed essi trovarono che la ritenzione andata a misurare nel test finale, era migliore per gli elementi che erano stati testati nel test di cued-recall rispetto agli elementi che non erano stati testati. Questo "effetti di test" si è ritenuto così poter essere la prova che le operazioni di recupero possono andare a modificare la traccia di memoria dell'elemento recuperato (Bjork, 1975, 1988; Cooper & Monk,

1976; Izawa, 1971, 1985a, 1985b; Bartling, 1980). Ma, nonostante ciò, per quanto riguarda il recupero in sé per sé, il cosiddetto retrieval, se effettivamente questo contribuisca agli effetti del test, questo ancora non si sa, perché non si sa effettivamente se un qualsivoglia effetto mnemonico di un test di memoria, non sia semplicemente dovuto alla ripresentazione del materiale target e non al recupero dalla memoria stessa (Skaggs, 1920). Critica che si applica a quegli studi che hanno dimostrato che ricevere un test di memoria porta ad una ritenzione successiva migliore rispetto alla performance che si otterrebbe, semplicemente avendo una quantità equivalente di tempo libero (Bartlett, 1977; Bartlett & Tulving, 1974; Darley & Murdock, 1971; Madigan & McCabe, 1971; McDaniel, Kowitz, & Dunay, 1989; Modigliani, 1976; Runquist, 1983, 1986; Young, 1971).

1.3.2 Materiale in prosa, saggi, test a risposta breve o scelta multipla

Propriamente l'effetto del test (o pratica di recupero) è ciò che ha portato a capire che la memoria a lungo termine è spesso aumentata, migliorata, quando una parte del periodo di apprendimento è dedicata al recupero delle informazioni da ricordare. Questo tipo di test non ha come obiettivo quello di poter **valutare la preparazione** di uno studente come la definizione a cui più comunemente si pensa quando si parla di test, quanto piuttosto quello di **migliorare l'apprendimento e la memorizzazione** delle informazioni da studiare. Esperimento cardine di questa letteratura sul Testing Effect è sicuramente quello di Roediger e Karpicke del 2006 nel quale i ricercatori andarono a testare la prestazione di alcuni studenti circa l'apprendimento di brani ricchi di informazioni, dividendo il gruppo dei soggetti in un gruppo a cui veniva somministrato il test, subito dopo aver letto i brani (condizione "studio/test"), ed un

secondo gruppo a cui veniva fatto leggere nuovamente il brano (condizione “studio/studio”). I due gruppi sono stati sottoposti ad altre 3 prove di valutazione: dopo 5 minuti, dopo 2 e 7 giorni.

I risultati hanno evidenziato come, a distanza di tempo (nelle prove a 2 e a 7 giorni di distanza), il gruppo che riportava una performance migliore fosse quello sottoposto a studio/test. Questo è propriamente il Testing Effect: sottoporsi ad un compito di richiamo piuttosto che rileggere il materiale, produce risultati migliori nella performance di memoria. (Carpenter et al. 2008 per una definizione simile). È importante però che il test venga eseguito subito dopo l'apprendimento iniziale, altrimenti la quantità di informazioni ricordate attraverso una prova di valutazione successiva, sarà scarsa e il richiamo non avrà la meglio sull'oblio della memoria.

Come quando si vuole imparare a suonare uno strumento musicale o a giocare a tennis o a ballare si deve far molta pratica prima di riuscire nell'intento desiderato, così, quando si vuole acquisire una o più conoscenze da un libro di testo si devono mettere in atto delle strategie per l'apprendimento. Comunemente si pensa che il modo migliore per apprendere sia rileggere più volte il materiale, far delle mappe concettuali o riassumere ciò che si è letto ma in realtà, se si vuole essere in grado di recuperare quelle conoscenze dalla memoria anche in un periodo successivo rispetto a quello in cui si è studiato qualcosa, la pratica del recupero sembra essere il modo migliore per imparare. Infatti, la pratica del recupero è uno dei modi più efficaci per consolidare la nuova conoscenza, per quanto questo spesso venga sottovalutato dagli studenti e dagli stessi educatori e considerato invece un metodo più di misurazione, piuttosto che di apprendimento vero e proprio.

Se ci si basa sul modello classico della memoria come costituita da tre magazzini a sé stanti quali codifica, memorizzazione e recupero, (Köhler,1947) e (Melton, 1963), le tre fasi sono da considerarsi temporalmente separate e solo dopo aver immagazzinato le informazioni nella memoria a lungo termine alcune di queste vengono recuperate, ma questa mentalità trascura i processi interattivi alla base del processo di apprendimento (Gardiner 2007).

Quando codifichiamo e poi recuperiamo, il processo di recupero fa sì che la memoria si consolidi e rimanga più accessibile nel tempo; potremmo considerare il recupero come una ri-codifica, un "modificatore di memoria" (Bjork 1975).

Inoltre, il potere di questa pratica risiede nel fatto che richiamare alla mente un'informazione in modo attivo, risulta essere più efficace della semplice rilettura del materiale studiato; questo perché il soggetto ne fa un'elaborazione personale, costringendosi ad uno sforzo maggiore rispetto ad un impegno cognitivo inferiore e sicuramente rispetto ad un'elaborazione più passiva quale quella della semplice rilettura. Al momento del richiamo, il soggetto testato partirà da degli indizi che sono quelle informazioni che ha elaborato ed organizzato in memoria nella fase precedente e grazie anche ad un altro fenomeno, detto "effetto di generazione" (*Baddeley, Eysenck, Anderson, La memoria, 2016*), la traccia mnemonica tenderà a rafforzarsi quando una data informazione viene rievocata autonomamente.

Ancora, in uno studio del 2013 di Meyer, A. N. D., & Logan, J. M. intitolato *Taking the testing effect beyond the college freshman: Benefits for lifelong learning*, gli autori hanno esaminato gli effetti dei test su 60 studenti universitari di età compresa tra 18 e 25 anni, 60 adulti di età compresa tra 18 e 25 anni e 60 individui di mezza età e anziani di età

compresa tra 55 e 65 anni con test immediati e con ritardo (2 giorni) su del materiale più complesso. I materiali di studio erano quattro articoli su armadilli, buchi neri, cuori umani e tsunami, argomenti di cui gli studenti conoscevano poco (verificato nei test pilota) (National Geographic, 2012a, 2012b, 2012c, 2012d).

Si è visto, in particolare, un miglioramento della memoria a lungo termine ottenendo risultati migliori in un test ritardato soprattutto se precedentemente testati sullo stesso materiale, rispetto ad una condizione di semplice restudying. Ristudiare infatti, spesso si traduce solo come un miglioramento della ritenzione ma a breve termine. Tutti i gruppi hanno beneficiato in modo simile dei test in entrambi i ritardi, il che implica che i test possono essere uno strumento utile di apprendimento permanente, indipendentemente dall'età dei soggetti e, si è visto emergere con diversi tipi di materiali: dai test vero/falso ai saggi, dalle coppie di numeri associati a parole a testi in prosa e quindi a materiale molto più rilevante dal punto di vista educativo.

1.3.3 Test con materiale più complesso

Definiamo i compiti di apprendimento “complessi” come quei compiti che hanno un alto grado di interattività tra gli elementi che li compongono (Sweller 2010; Sweller et al. 2011). Quando si imparano nuove informazioni che hanno un basso livello di interattività tra di loro, ogni elemento può essere appreso senza far riferimento per forza ad altri elementi appartenenti al compito. I materiali complessi invece, come testi didattici su fenomeni scientifici o sistemi meccanici o lo spartito che compone una melodia, sono tipicamente ad alta interattività di elementi,

correlati tra loro e che devono quindi essere elaborati contemporaneamente nella memoria di lavoro (Sweller et al. 2011).

Usare l'interattività degli elementi come misura della complessità del materiale non solo tiene conto della complessa struttura delle informazioni prese in esame ma anche, della conoscenza degli individui che vi si avvicinano: per qualcuno sarà più difficile e per qualcun altro più facile; perciò, il livello di complessità dipenderà anche dalle competenze del soggetto esposto al materiale (dello studente).

Presentiamo quindi di seguito alcuni esempi di studi che hanno suggerito come i test possano essere vantaggiosi anche per questi materiali ritenuti più complessi. Si è andato ad indagare l'effetto del testing rispetto ad una condizione di controllo di restudy su un test finale ritardato (avvenuto dopo qualche giorno o 1 settimana) dopo lo studio di Roediger e Karpicke (2006a) e i risultati raccolti sono contrastanti. Questo perché è stato dimostrato che l'effetto del test spesso diventa evidente solo con un ritardo di un paio di giorni o più, invece per periodi di tempo più brevi il restudy risulta essere talvolta anche più efficace (ad esempio, Roediger e Karpicke 2006a; Wheeler et al. 2003). Procedere nella ricerca con l'utilizzo di materiale più complesso è di estrema importanza per capire la reale utilità della pratica del recupero in ambito educativo. Per esempio, imparare elenchi di parole di una lingua straniera è necessario per realizzare compiti più complessi ed arrivare ad essere in grado di comunicare in quella lingua, capire la letteratura in quella lingua, o leggere e scrivere in quella lingua, che è l'obiettivo finale dell'educazione linguistica. Quindi la domanda è: l'effetto test si applicherebbe anche a compiti didatticamente rilevanti che sono più vicini agli obiettivi finali dell'istruzione, cioè all'apprendimento di materiali più complessi?

Van Gog et al. (2015) presentano quattro esperimenti sugli effetti del test quando si acquisiscono abilità di problem-solving attraverso lo studio di esempi. Hanno studiato gli effetti di diversi tipi di test come il richiamo libero dell'esempio, la risoluzione di un problema identico all'esempio presentato e la risoluzione di un problema isomorfo sulle prestazioni immediate e anche ritardate su un test finale di problem-solving con un altro esempio e, hanno visto come, in nessuno degli esperimenti emergeva un effetto test. Ma qui si potrebbe sostenere che la mancanza di un effetto test derivi dalla natura procedurale dei compiti di problem-solving. Mentre le coppie di parole o i fatti si basano sulla memoria dichiarativa, l'acquisizione di un'abilità di problem-solving richiede il passaggio dalla codifica dichiarativa iniziale alla memoria procedurale (ad esempio, Anderson 1982). Anche De Jonge et al. (2015) conducono diversi esperimenti per indagare l'effetto test usando del materiale diverso però; nel loro primo esperimento fanno imparare a degli studenti universitari un testo scientifico complesso di circa 1000 parole sui buchi neri. In una prima versione dell'esperimento, in cui ai partecipanti viene presentato per 15 minuti il testo proiettato su uno schermo una frase alla volta, poi viene loro detto di completare delle frasi con delle parti mancanti (condizione test) o continuare a studiare il materiale normalmente (condizione studio) e poi viene detto loro di completare un test finale immediato o ritardato di 1 settimana, i risultati dell'esperimento non hanno mostrato alcun effetto test. Ma in un secondo loro esperimento, in cui hanno usato lo stesso materiale ma le cui frasi erano state presentate in ordine rimescolato, un effetto test emerge, anche solo nel senso che il test sia andato a rallentare l'oblio rispetto al restudy. Anche ulteriori studi di Smith et al. (2016) evidenziano lo stesso fenomeno, ritornando all'idea originaria che sia propria la riduzione della complessità delle interazioni

tra gli elementi a far emergere tale effetto (Glover 1989). Nel suo studio “On imprinting through reading and reciting” Kühn (1914) riportò la prova che gli effetti benefici sulla memoria della recitazione (altro nome per la pratica di recupero o il test) rispetto al restudy, diminuivano all'aumentare della complessità del materiale didattico. Gates (1917) ha replicato i risultati di Kühn (1914) con un suo studio già citato in questo lavoro con l'utilizzo di sillabe senza senso e biografie e anche qui, l'effetto era più grande per le sillabe senza senso, benché l'interattività degli elementi delle biografie non fosse molto alta. In sintesi, questi studi sembrano suggerire che la complessità dei materiali di apprendimento può ridurre o addirittura eliminare l'effetto test.

Ma perché diminuisce con l'aumentare della complessità del materiale?

Sappiamo che i test possono influenzare la ritenzione sia direttamente attraverso l'atto della pratica del recupero stesso, che indirettamente influenzando lo studio futuro (grazie al feedback o altro) e anche se gli effetti diretti dei test sono stati stabiliti (vedi Roediger e Karpicke 2006b per una revisione), è possibile che il feedback sia spesso un prerequisito per trovare un effetto di test con materiali a bassa interattività di elementi. Con coppie di termini associati ad esempio, il test può rafforzare le tracce di memoria (direttamente quando il recupero ha successo o indirettamente tramite il feedback) ed il successo del tentativo di recupero non sembra essere importante, fino a quando viene fornito un feedback di risposta corretta (ad esempio, Kornell et al. 2009 e Richland et al. 2009). È importante però sottolineare come ogni studente può praticare il recupero da solo (con o senza un test vero e proprio); un test formale non è una parte fondamentale dell'equazione, né lo è alcuna risposta palese da parte dell'allievo. Cioè, la pratica di recupero silenziosa,

il feedback personale, è un efficace strumento di apprendimento al pari di altri (Putnam & Roediger 2013). Ancora, è stato dimostrato che i test di richiamo libero migliorano la ritenzione grazie al fatto che durante il richiamo avviene l'organizzazione delle informazioni (Gates 1917; Tulving 1962; Zaromb e Roediger 2010). Questa conclusione sembra essere in linea con i risultati di Masson e McDaniel (1981) che mostrano che la consapevolezza delle relazioni tra gli elementi durante lo studio favorisce anche le prestazioni in un test di ritenzione ritardata, presumibilmente a causa delle relazioni che si formano durante la codifica, come conseguenza del fatto che il recupero di un particolare elemento è facilitato dalla recuperabilità di altri elementi a cui è collegato. Per cui, questa funzione può essere obsoleta quando si apprendono informazioni complesse in cui gli elementi sono già altamente correlati (vedi anche Gates 1917).

In conclusione, si può dire che l'effetto testing si riscontra più su un materiale di minor complessità, che permetta alla memoria di elaborare le informazioni in modo da creare collegamenti tra gli elementi durante la codifica; ma anche il formato del test, il successo del recupero sul test iniziale o il feedback quando il recupero non ha successo, la spaziatura del test (Roediger et al. 2010), e anche i sentimenti di familiarità e motivazione, sono tutte possibili variabili che possono influire sulla possibilità di riscontrare o meno il testing effect; è un fenomeno complesso che sicuramente richiede ulteriori ricerche.

1.4 Testing online

1.4.1 Implementazione online di esperimenti in laboratorio

Prima di addentrarci nel vivo della nostra ricerca, è bene vedere come l'uso di internet per la ricerca scientifica, soprattutto in questi due anni di pandemia, ha dato la possibilità di poter continuare a raccogliere dati anche in situazioni in cui non era più possibile continuare a farlo direttamente nei laboratori di ricerca. Questo ha portato la ricerca non solo a non doversi fermare ma anche a doversi raffinare ed evolvere per poter raggiungere un livello di precisione ed affidabilità che in futuro potrà portare a benefici nella ricerca quali: grandi numeri dei campioni sperimentali, abbattimento dei costi dovuti alla preparazione di un esperimento in laboratorio. Non tutti gli esperimenti sono uguali però, e non è sempre facile tener a bada le variabili che potrebbero intaccare i risultati dei soggetti; In questo momento, la maggior parte degli studi mediati da Internet sembra essere basata su questionari, con test di personalità e sondaggi di vario genere che sono i più comuni e in piccola parte esperimenti sui tempi di reazione, sebbene Krantz, Ballard & Scher (1997), Schiano (1997) e Stern & Faber (1997) descrivono implementazioni elettroniche di altri paradigmi di ricerca.

Internet ha cambiato il modo in cui gli psicologi collaborano, raccolgono dati e i loro risultati per poi diffonderli. Internet sta cambiando il processo della ricerca empirica, portando ad opportunità e sfide. Internet offre enormi opportunità. Riduce molti dei costi della raccolta di dati sul comportamento umano, consentendo ai ricercatori di eseguire esperimenti online che coinvolgono migliaia di soggetti con costi molti ridotti rispetto

a quelli da sostenere per la costruzione di un paradigma sperimentale in laboratorio.

1.4.2 PRO e CONTRO della ricerca online

Pro:

- i costi di raccolta dei dati (L'uso di Internet riduce il costo del reclutamento di campioni ampi) e automatizzazione
- vedere interessanti fenomeni psicologici che non esistono negli ambienti tradizionali o sono lì difficili da studiare
- invisibilità dello sperimentatore, i ricercatori contaminano meno il setting sperimentale a differenza di un esperimento faccia a faccia
- i ricercatori hanno utilizzato Internet come fonte di dati sulle preferenze e le scelte individuali (Montgomery, 2001), sulla pigrizia sociale e sull'altruismo (Adar & Huberman, 2000) e sui modelli di amicizia (Adamic & Adar, 2001)

Contro:

- Internet solleva preoccupazioni sulla qualità dei dati e generalizzabilità dei dati e sul trattamento dei soggetti di ricerca
- I ricercatori spesso perdono il controllo sul contesto in cui i dati vengono acquisiti quando i soggetti partecipano a esperimenti online. Garantire il consenso informato, spiegare le istruzioni e condurre efficaci debriefing online può essere più difficile rispetto alle tradizionali impostazioni di laboratorio
- La costruzione di un paradigma sperimentale online richiede anche la conoscenza di software e relative competenze di programmazione

- Internet fornisce una nuova sede anche per la ricerca a lungo termine sui gruppi. Bruckman (1999).

Si guarda, ancora oggi, alla ricerca psicologica su internet come una pratica con cui dover procedere in modo cauto, giudicando spesso le tecniche di raccolta dati su Internet come “difettose” (Peterson, 2003) quando in realtà, non è così. Spesso le preoccupazioni che sono state mosse nei confronti della raccolta dati su Internet riguardano il fatto che i campioni di Internet non siano demograficamente diversificati (ad es. Krantz & Dalal, 2000), quando, sebbene i campioni di Internet non siano rappresentativi della popolazione in generale, sono generalmente più diversificati rispetto ai campioni pubblicati in una rivista di psicologia altamente selettiva. In particolare, i campioni Internet sono più rappresentativi dei campioni tradizionali per quanto riguarda il genere, lo stato socioeconomico, la posizione geografica e l'età e sono rappresentativi quanto i campioni tradizionali rispetto alla razza (D. Gosling, S. V. S. S. Oliver P. John, 2004). Altra critica mossa nei confronti del campione è che questo potesse essere composto per lo più da disadattati, o persone depresse, come nel caso degli Hikikomori (Mendenhall, 2001). Ma questa è una concezione obsoleta della visione di internet risalente agli inizi del fenomeno planetario, quando si pensava che internet potesse allontanare le persone dalla realtà (Kraut et al., 1998). E anche qui, si è visto che gli utenti di Internet non differiscono dai non utenti sui marcatori di adattamento e depressione. Inoltre, spesso si pensa che i partecipanti a Internet non siano motivati (ad es. Buchanan, 2000) come accade per i partecipanti durante l'esecuzione di un esperimento in loco ma non è così, anche con i metodi online possono essere usati dei feedback o dei modi per mantenere alta la motivazione dei partecipanti.

Ancora, le prove di esperimenti svolti fin ora, suggeriscono che i risultati basati su Internet sono coerenti con i risultati basati sui metodi tradizionali (ad esempio, sull'autostima, sulla personalità), ma sono necessari più dati ed ulteriori studi (John & Srivastava, 1999). Sicuramente sono state trovate alcune differenze quando veniva cambiato il formato di presentazione del materiale, e questo può avere gravi conseguenze negative per alcuni esperimenti (ad esempio con esperimenti su sensazioni e percezioni), ma per ricerche sui questionari, ad esempio, gli effetti della presentazione non sembrano compromettere la qualità dei dati. La natura anonima di Internet, inoltre, consente alle persone di partecipare in modo frivolo o con intenti dannosi e anche se non in modo volontario, i soggetti online possono semplicemente investire meno tempo ed energie nel compito di ricerca rispetto a quelli coinvolti in un sondaggio telefonico o in un esperimento di laboratorio. Si riportano infatti tassi di abbandono sostanzialmente più elevati nella conduzione di esperimenti online rispetto a quelli osservati durante ricerche simili in laboratorio (Williams, Cheung e Choi, 2000; Williams et al., 2002).

1.4.3 Sicurezza e standard per la protezione

Si ritiene che la ricerca online non rappresenti un rischio maggiore per i soggetti umani rispetto a ricerche comparabili condotte con altri mezzi, ma cambia ovviamente la natura dei rischi e la capacità di riuscire a valutarli.

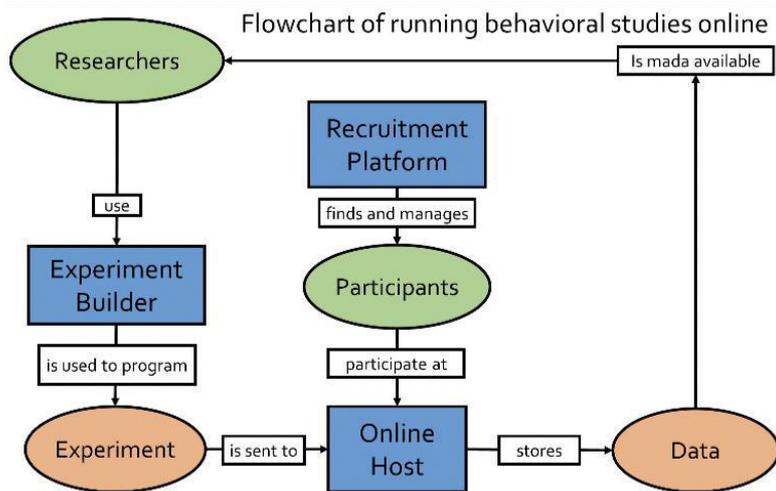
I concetti che stanno alla base della regolamentazione federale per la protezione dei soggetti umani cambiano o diventano ambigui quando la ricerca viene condotta online e nel Rapporto Belmont (Commissione nazionale per la protezione dei soggetti umani della ricerca biomedica e

comportamentale, 1979), questi principi sono stati formalizzati. Il regolamento fissa gli standard per la valutazione del grado di rischio per i soggetti umani e dei trade-off tra rischio e beneficio; per stabilire e documentare il consenso volontario e informato prima che le persone partecipino alla ricerca, e per il trattamento dei minori e di altre popolazioni vulnerabili. Il rischio maggiore associato alla ricerca online riguarda le violazioni della riservatezza, in cui vengono divulgate informazioni private e identificabili al di fuori del contesto di ricerca, ma, nel caso del sondaggio online e della ricerca sperimentale, il ricercatore può ridurre questo rischio non chiedendo esplicitamente informazioni identificative o registrando questi dati personali separatamente dai dati della ricerca (Tourangeau, Couper e Steiger, 2003).

In generale, i sondaggi, gli esperimenti o le osservazioni online non sono più rischiosi delle loro controparti offline. Per alcuni aspetti anzi possono risultare meno rischiosi perché il fatto di svolgere l'esperimento online, rende ancora più facile per i soggetti interrompere nel momento in cui si sentono a disagio. Questa libertà di ritirarsi non è un vantaggio banale, date le forti pressioni a continuare negli studi faccia a faccia (ad esempio, Milgram, 1963). Inoltre, alcuni passaggi di un esperimento o lo stesso feedback possono indurre i soggetti a riflettere su esperienze spiacevoli o ad apprendere qualcosa di spiacevole su sé stessi (ad esempio, Nosek et al., 2002b) e anche per gli esperimenti online si mette in atto un'analisi costi-benefici non diversa da quella della ricerca medica o psicologica tradizionale, in cui i rischi possono essere giustificati se i potenziali benefici della ricerca sono sufficientemente sostanziali. Inoltre, le linee guida etiche dell'American Psychological Association (2002) richiedono il debriefing dei soggetti, fornendo una spiegazione della natura dei risultati e delle conclusioni della ricerca anche per coloro che

abbandonano prima di completare la ricerca (Nosek, Banaji e Greenwald, 2002).

1.4.4 Costruzione di un esperimento online



Schema conduzione esperimenti online: gli esperimenti vengono creati con un generatore di esperimenti; i file sperimentali compilati vengono caricati su un host online, che genera un collegamento, rendendo lo studio accessibile online. I partecipanti vengono reclutati attraverso piattaforme di reclutamento e accedono agli esperimenti online sull'host. I dati sono archiviati sul server dell'host.

Analogamente agli studi progettati per essere svolti in laboratorio, la prima cosa da fare per la messa online di un esperimento è la sua programmazione e, se per la programmazione sperimentale di laboratorio si usano librerie Java, Python e MATLAB, per gli esperimenti online il linguaggio è invece il Javascript (JS). Le soluzioni attuali per la generazione sperimentale online spesso forniscono un'interfaccia utente grafica (GUI), che consente di semplificare i passaggi della programmazione, ma un buon ambiente sperimentale dovrebbe fornire la possibilità di estendere questi componenti modulari con script e soluzioni basate su codice. Sicuramente OpenSesame e PsychoPy Builder sono tra

gli ambienti più popolari che consentono una ristrutturazione di un paradigma sperimentale nella sua controparte online in cui bisogna però riscrivere tutte le sezioni in cui è stato utilizzato lo scripting in linguaggio Javascript. (Fig.2)

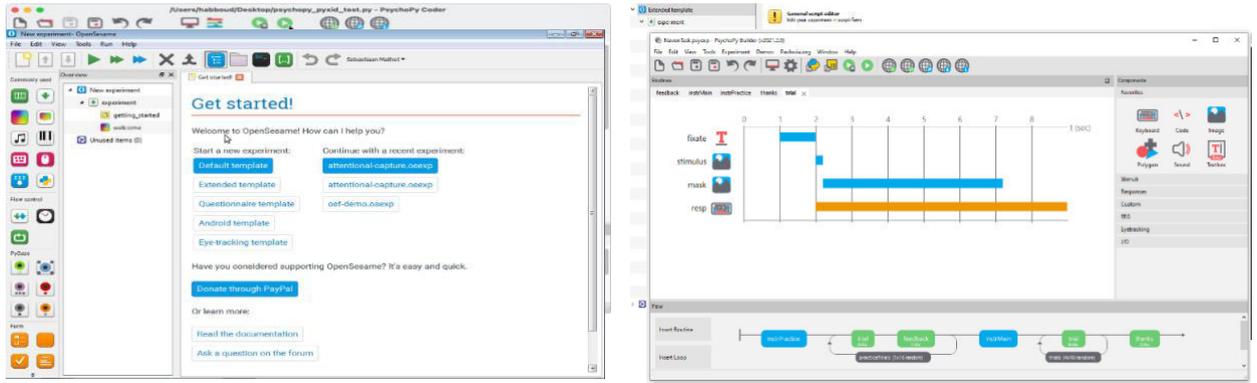


Fig.2 Esempio delle interfacce grafiche con cui si presentano OpenSesame (a sinistra) e PsychoPy (a destra) con le loro relative schermate di script.

Fortunatamente, Python e Javascript differiscono solo in termini di sintassi e non di logica di programmazione per cui la ristrutturazione di un paradigma sperimentale si rivela essere relativamente facile. Ci sono inoltre molti altri costruttori di esperimenti disponibili: Gorilla, Inquisit Web, LabVanced e Testable, jsPsych e altri i cui vantaggi e svantaggi dovrebbero essere valutati in laboratorio a seconda delle esigenze individuali del ricercatore. In generale, però, operando quasi tutti su linguaggio Javascript, offrono una flessibilità simile fatte alcune eccezioni.

Per gli studi online, l'esperimento deve essere reso disponibile per la distribuzione online ospitandolo su un server (di solito un host). Alcuni laboratori con molta esperienza nella sperimentazione online ospitano questi studi sui propri server. Ciò comporta i vantaggi di bassi costi di

manutenzione, pieno controllo e flessibilità. Tra gli svantaggi, richiede una certa esperienza per l'installazione e la manutenzione continua degli stessi però. L'alternativa più fattibile per coloro che hanno meno esperienza o meno risorse, sono i provider di hosting centralizzato. Qui, l'hosting e la gestione dello studio sono un servizio che viene fornito a pagamento e che prevede una gestione semplificata (per lo sperimentatore) dei partecipanti e dell'archiviazione automatizzata dei dati.

Ad esempio, uno dei servizi di hosting più interessanti è Pavlovia, uno spazio in cui i ricercatori in scienze comportamentali possono condividere ed eseguire esperimenti online. Esso offre una facile integrazione con strumenti di reclutamento (mette a disposizione piattaforme come Sona e Prolific per reclutare partecipanti) e una piattaforma GitLab in cui gli sperimentatori possono condividere il loro codice completo. Ma bisogna tener conto della compatibilità esperimento-server. Nessun host supporta tutte le build sperimentali e nessuna build sperimentale è compatibile con tutti gli host. Pertanto, una decisione dovrebbe sempre essere presa a livello dell'ecosistema dell'esperimento nel suo complesso.

Il vantaggio principale dell'esecuzione di studi sperimentali online risiede nella sua efficienza. Come si è detto nei paragrafi precedenti, è possibile in questo modo raccogliere le risposte di centinaia di partecipanti in poche ore. Grazie alla possibilità di campionamento in tutto il mondo, la raccolta dei dati può essere completata letteralmente nell'arco della notte. Una volta che l'esperimento è stato creato e risulta accessibile online attraverso un certo tipo di collegamento, i partecipanti possono essere reclutati. A causa del numero più elevato di partecipanti rispetto alla maggior parte degli studi di laboratorio, non è consigliabile gestire questo

processo manualmente quanto piuttosto si consiglia di utilizzare dei sistemi di gestione del pool di partecipanti, che offrono strumenti di automazione completi. Tuttavia, si consiglia di mantenere il proprio pool di partecipanti di dimensioni limitate per averne un maggior controllo. Inoltre, solo un numero limitato di partecipanti può essere reclutato dall'Università locale, tramite social media e mailing list (istituzionali). Ci sono molti fornitori che offrono servizi di reclutamento dei partecipanti e gestione dei pagamenti e degli strumenti di targeting come Amazon Mechanical Turk (MTurk), Prolific Academic e Qualtrics Panel e poiché è necessario solo che ci sia un collegamento allo studio, si integrano bene con i sistemi di gestione dello studio e con i costruttori di esperimenti precedentemente menzionati.

Riassumendo, ciò di cui hanno bisogno i ricercatori per mettere in atto una sperimentazione online, si può dire che i passaggi siano in linea di massima gli stessi che bisogna rispettare per uno studio in laboratorio: un esperimento programmato, un server per ospitare lo studio e una piattaforma di reclutamento per i partecipanti e per far questo ci sono soluzioni differenti. I provider di servizi integrati consentono di risparmiare tempo riducendo i problemi di compatibilità, fornendo assistenza ai clienti e riducendo il carico amministrativo; anche se, a volte possono mancare di trasparenza e flessibilità, o minima compatibilità tra i sistemi e generalmente risultano essere anche molto costosi. Le soluzioni invece meno costose, di solito non dispongono dell'assistenza diretta del cliente, ma forniscono forum e feedback della community, costi bassi o nulli (vedi Pavlovia). Per cui la scelta di quale ecosistema online utilizzare dovrebbe essere presa in base alle capacità e alle esigenze del laboratorio, ma poiché tutte le piattaforme sono basate su Javascript, la costruzione e la messa in atto di un esperimento online dovrebbe essere teoricamente

realizzabile su tutte le piattaforme, anche se a volte, gli esperimenti precedentemente programmati non possono più essere eseguiti anche mettendo in atto degli aggiustamenti nell'esperimento (e quindi nello script) ma è necessario ristrutturare completamente il progetto. Pertanto, i ricercatori, quando decidono come trasferire gli esperimenti nel mondo online, non dovrebbero considerare solo ciò che il provider offre, ma anche come adattare la loro ricerca al nuovo ambiente. Si sono documentate buone esperienze con OSWeb (per la costruzione) combinato con JATOS (per l'hosting) e PsychoPy (per la costruzione) con Pavlovia (per l'hosting) e Prolific (per il reclutamento) (M. Sauter, D. Draschkow, W. Mack, 2020) e contemporaneamente, i ricercatori più esperti sconsigliano di avviare studi sperimentali su server web autogestiti senza l'ausilio di un sistema di gestione dello studio come JATOS poiché si potrebbe perdere il controllo sull'archiviazione dei dati ma anche l'assegnazione dei codici ai partecipanti (perdendo così la sicurezza che i partecipanti non partecipino più di una volta). Noi stessi, per il nostro esperimento, abbiamo utilizzato una combinazione di Jatos, PsychoPy e Pavlovia, mentre il reclutamento dei soggetti è avvenuto tramite archivi universitari e il passaparola.

Riprendendo quello che si è già affrontato nei paragrafi precedenti, ovvero la preoccupazione per la qualità dei dati, si può affermare che mentre le preoccupazioni più ovvie come la motivazione, i fattori distraenti e i tempi di reazione e il timing degli stimoli possono essere affrontate con una progettazione appropriata, una valutazione più attenta bisogna che si abbia nei confronti dei casi di abbandono. Si può immaginare che i tempi di presentazione degli stimoli o i tempi di risposta non siano affidabili a causa delle variazioni della velocità di Internet o delle impostazioni di visualizzazione durante l'esperimento; tuttavia,

quasi tutte le soluzioni online funzionano scaricando (pre-buffering) l'intero esperimento sulla macchina del partecipante. E ancora, l'utilizzo di una virtual chinrest per calcolare l'esatta distanza da mantenere rispetto allo schermo (Q. Li, S. Jun Joo, J. D. Yeatman e K. Reinecke 2020, vedi appendice), e il fatto che le moderne frequenze di aggiornamento dello schermo sono quasi esclusivamente impostate su 60 Hz (standard di fatto), rende alcune variabili degli studi online un po' più prevedibili. Due recenti ampi studi hanno studiato la precisione temporale nella presentazione dello stimolo di diverse soluzioni online e offline; ed il confronto online ha riscontrato una buona precisione complessiva per Gorilla, jsPsych e PsychoJS e lab.js con prestazioni però particolarmente scarse utilizzando il browser Safari e Mac OS X. Il confronto offline con PsychoPy e Opensesame ha raggiunto comunque una precisione compresa tra 1 ms e 4 ms. Uno studio del 2017 di Semmelmann, K.; Weigelt, S. (Online psychophysics: Reaction time effects in cognitive experiments) che ha esaminato i tempi di risposta, ad esempio, ha trovato un offset temporale additivo di 87 ms (simile in tutte le condizioni) nelle registrazioni online rispetto agli studi di laboratorio, riproducendo al contempo tutti gli effetti attesi in vari compiti (stroop, flanker, visual search, attentional blink).

Per quanto riguarda i tassi di abbandono, a differenza degli studi di laboratorio, si è visto che negli studi online i partecipanti possono abbandonare gli studi fino al 69%. In un'analisi di abbandono scolastico di 88 studi locali, Zhou e Fishbach è stato riscontrato che il 20% aveva un tasso di abbandono scolastico superiore al 30%. In modo allarmante, gli autori degli studi analizzati non erano a conoscenza di questi abbandoni e questi tassi di abbandono possono interagire con la condizione sperimentale.

Oltre ai tempi e alla preoccupazione sull'abbandono da parte dei partecipanti, potrebbero esserci preoccupazioni anche sul fatto che i partecipanti possano essere meno impegnati quando da casa non vengono osservati direttamente dallo sperimentatore. Diversi studi hanno dimostrato che non si trova necessariamente una ridotta attenzione al compito (Clifford, S.; Jerit, J. 2014, Hauser, D.J.; Schwarz, 2015) e la qualità dei dati è paragonabile agli studi di laboratorio. In uno studio recente, i partecipanti sono stati sottoposti a diversi controlli sull'attenzione e non c'era differenza tra i partecipanti al laboratorio e i partecipanti online in nessuna delle misure indagate. Tuttavia, lo studio ha mostrato che i partecipanti online avevano una maggiore distrazione auto-riferita (uso del telefono cellulare, parlare con un'altra persona, ecc.). In ogni caso, gli sperimentatori dovrebbero adattare il loro esperimento per tenere conto della diversità del campione e delle motivazioni dei partecipanti e del setting di fronte al quale si trovano. Alcuni autori consigliano che questo controllo possa essere realizzato fornendo accuratezza come feedback dopo ogni prova e fornendo valutazioni sintetiche delle prestazioni dopo blocchi di prove (Crump, M.J.C.; McDonnell, J.V.; Gureckis, 2013).

Le istruzioni sperimentali per un esperimento online inoltre devono necessariamente essere abbastanza facili ed intuitive, in modo da essere comprese da un campione diversificato. Gli sperimentatori devono essere più accurati durante la creazione di queste istruzioni sperimentali e accertarsi che possano essere comprese anche senza spiegazioni verbali ulteriori. È stato visto come un'istruzione pittorica (visiva) porti a meno incomprensioni o addirittura meno abbandoni rispetto a una singola pagina di testo ed è consigliabile che le istruzioni siano forzate a rimanere sullo schermo per un certo periodo di tempo prima che sia consentita la

continuazione. È necessario adottare misure per rilevare e scoraggiare le scarse prestazioni, ovvero la partecipazione "falsa" ed infine, si consiglia una durata minore per gli studi online rispetto al loro corrispettivo offline in laboratorio.

Per quanto riguarda l'interazione tra sperimentatore e partecipante, questa risulta essere relativamente indiretta negli esperimenti online. È consigliabile quindi affermare esplicitamente la rilevanza della ricerca; è stato visto come questo porti a prestazioni migliori da parte dei partecipanti (Chandler, D.; Kapelner, 2013)

Per molti partecipanti provenienti dai servizi di reclutamento, la motivazione dominante per la partecipazione è un compenso monetario, anche se, la qualità dei dati non è necessariamente influenzata da incentivi monetari più elevati (Crump, M.J.C.; McDonnell, 2013). Tuttavia, si è evidenziato che potevano raccogliere i dati più rapidamente e avevano meno abbandoni quando il pagamento era più alto.

Infine, gli studi sperimentali online dovrebbero essere più brevi rispetto agli esperimenti eseguiti in laboratorio. In uno studio si è chiesto a dei partecipanti in quale momento avrebbero interrotto un esperimento online che offriva un salario minimo. La maggior parte degli intervistati ha affermato "dopo 15 minuti" (44%), seguito da "dopo 30 minuti" (35%), "dopo 45 minuti" (10%) e "dopo 60 minuti o mai più" (12%) (Appinio GmbH. 2020).

In conclusione, si può quindi affermare che, per un certo sottoinsieme di indagini, gli studi online che vengono sviluppati con tutte le accortezze del caso, hanno un potenziale enorme nella ricerca odierna. Molti dei fattori di rumore possono essere ridotti al minimo con un lavoro preparatorio intelligente. È certo che i ricercatori debbano scegliere con saggezza e attenzione il software da utilizzare e in alcuni casi l'host, in

base alle risorse del proprio laboratorio e ai requisiti da dover rispettare secondo il dominio di studio preso in esame.

2. LA RICERCA

2.1 Obiettivo dello studio

Il presente elaborato si occupa di indagare il fenomeno definito in letteratura *testing effect*, nel dominio della memoria di riconoscimento di scene naturali. Lo scopo generale dello studio è quello di indagare il contributo del test di memoria sul consolidamento a lungo termine, utilizzando immagini che variano nel contenuto e nel grado di nitidezza. Nello specifico, il presente lavoro di tesi si è occupato di replicare uno studio precedentemente svolto in laboratorio, introducendo la variante online, che ha previsto la raccolta dati attraverso la piattaforma Pavlovia.

Il paradigma sperimentale adottato a tale scopo è stato sviluppato in tre fasi. Una prima fase di Codifica, una fase di pratica che è propriamente il Testing e un'ultima fase di Test di riconoscimento *old-new* finale.

Le scene naturali utilizzate hanno subito una manipolazione del contenuto emozionale in modo da contrapporre a immagini emozionali (piacevoli e spiacevoli) immagini neutre e, inoltre, si è andato a manipolare il grado di nitidezza del materiale usato in fase testing (una quota di immagini presentate in modo nitido, una quota invece presentata degradata, sfocata e quindi povera di dettagli).

Come il tipico paradigma del *Testing Effect* prevede (Carpenter, 2006), tra il primo test e quello finale è stato interposto un intervallo temporale di una settimana. Per quanto riguarda il lavoro qui presentato, sia nella fase di Testing che in quella di riconoscimento *old-new* finale sono

stati registrati i soli dati comportamentali e, si è andati a vedere se la manipolazione online risulta affidabile nella raccolta dati quanto un esperimento svolto in laboratorio o se da remoto intervengono variabili che ne riducono l'attendibilità. L'utilizzo di questa modalità online, se affidabile, risulterebbe utile nell'introdurre nuove manipolazioni nel paradigma per studi futuri.

Poiché si è già visto in precedenza in diversi studi che una manipolazione delle informazioni (attraverso un test) produce un beneficio maggiore nel ricordo dell'intero materiale rispetto al ristudio, ci si aspetta di confermare questo dato anche attraverso l'esperimento online. Ci si aspetta, come nello studio precedente, un vantaggio delle immagini integre rispetto a tutte le altre e si vuol chiarire quale sia il meccanismo in gioco nel consolidamento di queste immagini. Ci si chiede se queste immagini integre vengano ricordate meglio rispetto alle altre in fase finale perché ricordate maggiormente anche in fase testing oppure se il contributo del testing sia diverso e risulti più efficace con questa quota di immagini più ricche di dettagli. Per andare ad indagare il reale contributo del Testing, si è fatto un confronto tra la fase Testing e la fase di Test finale, analisi che non si era potuta eseguire nell'esperimento precedente per mancanza di dati del testing. Infine, ci si aspetta di trovare un generale vantaggio mnestico per gli stimoli con valenza emozionale indipendentemente dalla condizione testing.

2.2 Materiali e metodi

2.2.1 Soggetti

L'esperimento è stato condotto online nel pieno rispetto delle disposizioni a seguito dell'emergenza sanitaria Covid-19. Allo studio

hanno preso parte 14 soggetti, di cui 6 maschi e 8 femmine, con un'età compresa tra i 18 e i 27 anni (media = 22, deviazione standard = 2.5). Inizialmente il campione raccolto ammontava a 66 soggetti, la maggior parte esclusi per problemi legati al proprio sistema operativo usato o altro (vedi Appendice D-Totale soggetti).

Prima di accedere all'esperimento vero e proprio, il soggetto si prestava ad un compito di Virtual Chinrest, per poter assumere la giusta distanza dallo schermo prima di iniziare il task sperimentale. Questo compito rientra in una delle misure prese per eliminare possibili *bias* di valutazione interindividuale (vedi Appendice B-Virtual Chinrest).

Dopo il calcolo della distanza dallo schermo, tutti i soggetti sperimentali hanno visionato e firmato il consenso informato in formato digitale prima di procedere con qualunque prova dell'esperimento. Poi eseguivano una prima parte di anamnesi, relativa al proprio stato civile, professione, ore di sonno nella notte precedente e stato di salute mentale e fisica (in cui dichiaravano di avere una vista normale, o comunque corretta alla normalità tramite lenti o occhiali da vista).

2.2.2 Stimoli

Sono state utilizzate in tutto 494 immagini selezionate dallo IAPS (International Affective Picture System) e da risorse WEB, di cui 14 buffer ovvero, immagini irrilevanti poste all'inizio e alla fine del test come stimoli di controllo. Di queste immagini, 180 sono state presentate durante la fase di codifica iniziale, successivamente solo 120 sono state riproposte nella fase intermedia di Testing, insieme ad altre 120 immagini nuove, cosiddette *filler*, mentre le rimanenti 60 sono state riviste direttamente al test finale. Nella fase finale, al test di riconoscimento *old-new*, sono state

mostrate 364 immagini, di cui 4 immagini buffer, 180 immagini presentate in codifica e 180 nuove (vedi Tabella 2.1). Di questi 480 stimoli:

- 240 sono immagini neutre, facenti parte di 10 categorie differenti (24 stimoli a categoria in totale): animali, ambienti esterni, ambienti interni, volti, oggetti, strumenti, ambienti esterni con figure umane, ambienti interni con figure umane, mezzi di trasporto, stanze di casa.
- 120 sono immagini con valenza piacevole, appartenenti a 5 categorie differenti (24 stimoli a categoria in totale): bambini, scene erotiche, immagini romantiche, animali domestici e sport estremi.
- 120 sono immagini con valenza spiacevole, appartenenti a 5 categorie differenti (24 stimoli a categoria in totale): paura, disgusto, minaccia, mutilazioni e immagini ritraenti tristezza.

2.2.3 Paradigma

Lo studio si offre di indagare il fenomeno del *Testing Effect*, nel dominio delle immagini, eseguendo però l'esperimento da remoto (online). Lo studio è stato sottoposto a 66 persone ma dalle analisi sono stati esclusi la maggior parte dei soggetti per problemi legati al caricamento durante le prove svolte dal proprio laptop o per aver interposto una distanza maggiore di 1 settimana tra la prima prova e il test finale di riconoscimento old-new (vedi Appendice D). Il paradigma sperimentale è stato sviluppato in tre fasi: una prima fase di Codifica, una fase di Test (Testing) e un'ultima fase di Test di Memoria di riconoscimento finale. Nella fase di Codifica, i partecipanti sostenevano

un compito di memoria incidentale, ossia erano inconsapevoli che successivamente sarebbero stati testati in un compito di riconoscimento.

Gli stimoli sono fotografie di scene naturali, con contenuti eterogenei, anche di natura emozionale (scene romantiche, erotiche, incidenti, violenza, ecc..). Tutte le immagini sono materiale tratto dallo IAPS (International Affective Picture System) e da risorse WEB (Bradley, M. M., & Lang, P. J. 2007). Si tratta di set di immagini standardizzate sottoposte a rating per arousal e valenza. Visionavano delle immagini dinamiche che da degradate diventavano nitide e tornavano ad essere degradate (immagini gif). Il vincolo che abbiamo imposto alla randomizzazione è stato di non presentare più di tre immagini consecutive con la medesima valenza. L'unica istruzione data ai partecipanti era quella di osservare attentamente le stesse.



Figura 2.1: Esempi di stimoli utilizzati nell'esperimento

Sono state presentate 190 immagini dinamiche in formato .gif (di cui 10 buffer, 5 all'inizio e 5 in chiusura alla sequenza) tramite il programma Psychopy.

CODIFICA	TESTING Old-New		1 settimana	TEST FINALE Old-New	
	180 immagini di cui	120 OLD di cui		120 NEW di cui	180 OLD di cui
90 neutre	30 integre	30 integre		90 neutre	90 neutre
	30 degradate	30 degradate			
45 piacevoli	15 integre	15 integre		45 piacevoli	45 piacevoli
	15 degradate	15 degradate			
45 spiacevoli	15 integre	15 integre		45 spiacevoli	45 spiacevoli
	15 degradate	15 degradate			

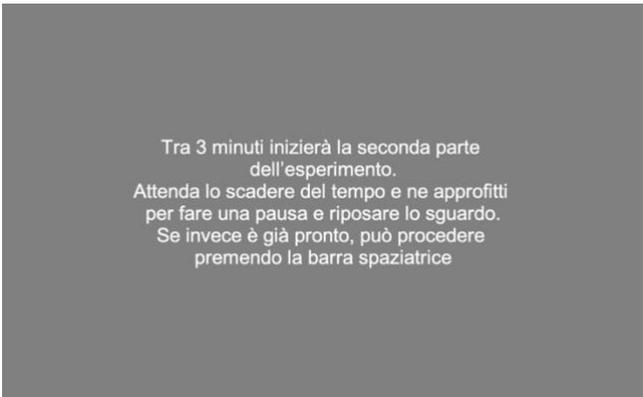
Tabella 2.1: Ripartizione degli stimoli: in Codifica sono state presentate 180 immagini, di queste solo 120 hanno preso parte al Testing, 60 presentate integre e 60 degradate. Infine, a distanza di una settimana sono state ripresentate le medesime scene presentate in codifica più 180 nuove. Il verde fa riferimento alle immagini neutre, il blu a quelle piacevoli e il rosso a quelle spiacevoli.

Tali stimoli sono stati acquisiti dallo IAPS (International Affective Picture System, cit) e da risorse WEB. Queste gif sono state prodotte tramite uno script di MATLAB 2015 (vedi Appendice D). Tra le 190 immagini sono presenti 10 buffer, che vengono presentate 5 all'inizio e 5 alla fine della sequenza per evitare effetti di posizionamento seriale; quindi, sono state eliminate successivamente dalle analisi.

In fase di codifica, ogni *trial* era una gif della risoluzione di 800×600 megapixel, che veniva proiettato per una durata di 5 secondi, con un ITI (Inter Trial Interval) composto da una fase di fissazione di 2 secondi in cui il soggetto visualizzava una croce centrale di colore bianco e una seconda fase di fissazione in cui la croce diventava verde, che il partecipante doveva premere, tramite mouse, per andare avanti e poter visualizzare il trial successivo (Figura 2.2).



Tale croce di fissazione di colore verde aveva una durata massima di 120 secondi. Avendo costruito il paradigma affinché potesse essere svolto online, in un luogo che non fosse un laboratorio, sono state introdotte diverse pause per ridurre il più possibile la disattenzione dei soggetti legata alla stanchezza. A tale proposito, dopo i *trial* 60, 120 e 190 è stata introdotta una pausa di 60 secondi, per permettere ai partecipanti di riposare; al termine della pausa fissa, iniziava un'altra pausa della durata massima di 180 secondi in cui, tramite una consegna, veniva detto al soggetto che avrebbe potuto procedere alla fase successiva premendo la barra spaziatrice se si fosse sentito già pronto, oppure avrebbe potuto attendere ancora fino al suo termine.



Tra 3 minuti inizierà la seconda parte
dell'esperimento.
Attenda lo scadere del tempo e ne approfitti
per fare una pausa e riposare lo sguardo.
Se invece è già pronto, può procedere
premendo la barra spaziatrice

Delle 180 immagini utilizzate ai fini dell'esperimento, 90 immagini presentano contenuto neutro e 90 presentano contenuto emozionale, divise a loro volta in 45 a contenuto piacevole e 45 spiacevole. Di questa quota soltanto 120 verranno utilizzate nella fase successiva e le rimanenti 60 verranno presentate di nuovo nel test finale a distanza di una settimana. La fase di Testing è composta da due blocchi, intervallate da una pausa di 3 minuti simile alla precedente (cioè dopo 60 secondi di fermo obbligatorio, potevano decidere quando ricominciare, se aspettare ancora fino allo scadere dei 3 minuti o riprendere prima). Una sezione ha previsto la presentazione di immagini integre e una di immagini degradate (Fig.2.3). Ai partecipanti è stato richiesto di compiere un compito di riconoscimento *old-new*, tramite la selezione di due tasti, "K" per indicare le immagini già viste e "M" per quelle nuove. È richiesto che la risposta di discriminazione venga espressa con la mano dominante.



Fig. 2.3: Esempi di *trial* mostrati nella fase 2, quella di Testing. I partecipanti dovevano svolgere un riconoscimento *old-new*, selezionando il tasto "k" per gli stimoli già visti e "m" per quelli nuovi.

Le immagini all'interno di ogni sequenza erano state bilanciate in modo tale che non ci fossero più di tre immagini nuove o vecchie consecutivamente, e non più di tre immagini con la stessa valenza. Per

ciascuna delle due condizioni, integre e degradate, sono state presentate 120 immagini, per un totale di 240 stimoli sperimentali 800×600 megapixel. Di queste 240 immagini, 120 sono esattamente le stesse presentate nella fase di codifica, di cui 60 sono state mostrate integre e 60 degradate, le restanti erano nuove che, per uniformità nel compito, sono state presentate 60 integre e 60 degradate. Oltre alla degradazione, è stata variata la valenza del contenuto. Dei 240 stimoli in questione, 120 erano di natura neutra e 120 emozionali, di cui 60 piacevoli e 60 spiacevoli, bilanciati per le condizioni integro e degradato. Il materiale è stato convertito in bitmap, mediante IRFANVIEW e presentato tramite E-PRIME® nell'esperimento tenuto in laboratorio e qui, presentato invece tramite PsychoPy e quindi Pavlovia (vedi paragrafo 1.3.4) per una durata di 1 secondo, dopo il quale appare una schermata che suggerisce al soggetto di premere il tasto “K” per un'immagine che pensano di aver visto e “M” per una che invece non credono di aver visto prima. Questa schermata non ha un termine, e si può procedere all'immagine successiva solo effettuando la risposta, in modo tale da non perdere le risposte nell'output. Segue un intervallo di 2 secondi prima di procedere all'item successivo.

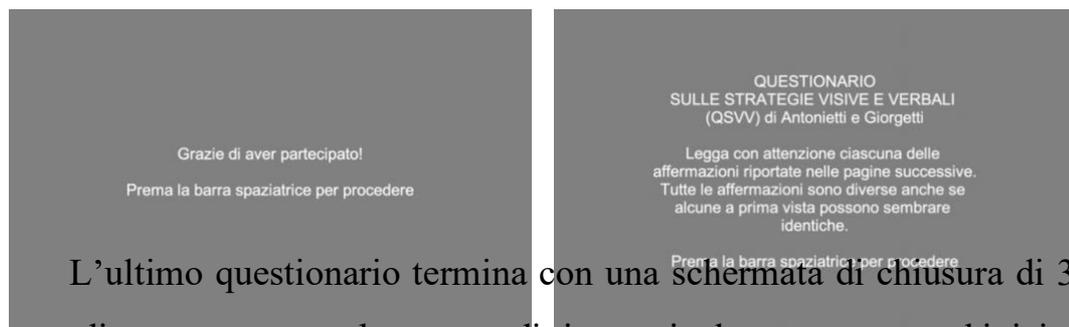
Nelle fasi di testing e di test di memoria finale i soggetti rivedono le stesse immagini viste in codifica insieme ad uno stesso numero di immagini mai viste prima, e dovranno indicare quali immagini sono nuove e quali invece già viste. Come il tipico paradigma del Testing Effect prevede (Carpenter, 2006), tra il primo test (fase testing) e quello finale, è previsto un intervallo temporale di una settimana. I partecipanti ricevono, pertanto, un primo link che li porta direttamente alla piattaforma su cui è caricato l'esperimento e un secondo link, a distanza di una settimana, con cui svolgere il secondo esperimento (il test di Memoria). Nella seconda parte dell'esperimento, saranno presentati dei questionari prima della

sessione sperimentale vera e propria. Al fine di evitare che vengano reclutati partecipanti particolarmente sensibili verso specifiche scene emotive (ferite), si somministra il questionario Fear Survey Schedule (FSS) per valutare eventuali fobie.

I tre questionari presentati (per comprendere quali costrutti misurano vedere paragrafo 2.2.5) sono:

- Fear Survey Schedule (FSS III) (Wolpe & Lang, 1964);
- Verbal-Visualizer questionnaire (VVQ) (Richardson, 1977);
- Questionario di strategie visive-verbali (QSVV) (Antonietti e Giorgetti, 1993);

A differenza di una versione cartacea precedentemente utilizzata in laboratorio, in questa versione online, i questionari vengono presentati all'interno della sessione sperimentale della seconda settimana, prima del test di memoria finale (old-new) ed ogni questionario viene separato dal successivo tramite una schermata finale del questionario corrente di 15 secondi, seguita dalla schermata iniziale del secondo questionario della durata di 20 secondi.



L'ultimo questionario termina con una schermata di chiusura di 30 secondi per permettere al soggetto di riposarsi ed essere pronto ad iniziare l'esperimento di memoria old-new finale vero e proprio.

Figura 2.4: esempio di fine primo questionario FSS III ed inizio secondo questionario QSVV

Nel test finale sono stati presentati 364 stimoli, di cui 4 buffer, 2 iniziali e 2 finali. Dei 360 stimoli presi in analisi, 180 erano quelli presentati durante la fase di codifica (120 presentate anche nel testing e 60 solo in codifica iniziale), mentre i restanti erano nuovi, per cui 180 erano

neutre, 90 piacevoli e 90 spiacevoli e tutte presentate nella loro versione integra, equamente divise per la condizione old e new. Tutte le immagini sono presentate per 1 secondo, dopo il quale appare una schermata che suggerisce al soggetto di premere il tasto “K” per un’immagine che pensa di aver visto e “M” per una che invece non crede di aver visto prima. Dopo la risposta da parte del soggetto appare una croce bianca per la durata di 2 secondi che anticipa l’immagine successiva. Dopo per prime 120 immagini c’è una pausa di 180 secondi durante la quale viene consigliato al soggetto di riposare lo sguardo; ci sono tre pause da 180 secondi ogni 120 immagini per il totale di 364 immagini presentate. Al termine dell’esperimento appare una schermata finale di 10 secondi durante la quale avviene il caricamento del file del soggetto.

2.2.4 Procedura

Il reclutamento dei partecipanti avviene tramite e-mail universitaria, contattando studenti dell’Ateneo di Parma e Bologna e non solo. I partecipanti che hanno dato l’adesione hanno ricevuto una mail con il link per l’esperimento, consenso informato (presente anche in formato cartaceo che si richiede comunque di stampare, firmare e inviare nuovamente) e relative istruzioni per eseguire il task.

I soggetti vengono avvisati del fatto che alla fine dell’esperimento potrebbero passare alcuni minuti per il caricamento dei risultati che verranno raccolti sia su Pavlovia (piattaforma online su cui vengono caricati i progetti Psychopy, vedi cap 1.3.4) sia sul server Jatos

(piattaforma su cui vengono creati i link collegati a Pavlovia, da inviare ai partecipanti) e per questo viene chiesto loro di attendere.

I partecipanti sono, inoltre, informati fin dall'inizio del fatto che, qualora la situazione sperimentale dovesse causare loro disagio, saranno liberi di interrompere l'esperimento in qualunque momento. L'esperimento durerà circa 60 minuti per sessione (pause escluse). Per riportare online lo stesso paradigma dell'esperimento eseguito in laboratorio, si sono dovute apportare alcune modifiche per assicurarsi che i soggetti eseguissero l'esperimento nel modo più corretto possibile ovvero, cercando di prevedere ed escludere tutte quelle distorsioni che potrebbero intervenire in situazioni meno controllate di rispetto al setting da laboratorio di ricerca.

Per caricare un esperimento online, per prima cosa bisogna decidere il programma da utilizzare per costruire il paradigma sperimentale. In un primo momento abbiamo iniziato la costruzione dello stesso su *Opensesame* per poi spostarci su *Psychopy* (vedi 1.3.4) poiché non tutti i programmi di *experiment building* permettono di esportare online il materiale creato. *Psychopy* permette di caricare in sicurezza il progetto creato in Pavlovia, piattaforma online, e condividerli pubblicamente attraverso l'accesso ad un link.

Si è voluto mantenere il paradigma quanto più simile a quello originario, per cui sono state create 8 sequenze che differivano per l'ordine di presentazione delle immagini, corrispondenti ad altrettanti link diversi. Il link veniva inviato al partecipante in una mail in cui, come si sarebbe fatto in laboratorio, il soggetto leggeva e firmava il consenso informato attraverso firma digitale e poi leggeva le istruzioni per poter procedere all'esperimento, accedendo al link. Prima di accedere all'esperimento vero e proprio il soggetto si prestava ad un compito di *virtual chinrest*, per

poter assumere la giusta distanza dallo schermo prima di iniziare il task sperimentale (Q. Li, S. Jun Joo, J. D. Yeatman e K. Reinecke 2020). Questo compito rientra in una delle misure prese per eliminare possibili bias di valutazione tra i vari soggetti. Dopo il calcolo della distanza dallo schermo, si procede con ulteriore consenso informato al quale il soggetto risponde acconsentendo o meno con una risposta dicotomica si/no e con alcune domande di anamnesi di livello generale (età/sexo/condizioni di vista). Conclusa questa fase iniziale, parte la fase di Codifica con la presentazione delle immagini formato .gif che il soggetto deve solo visionare. Al termine della fase di Codifica seguiva dopo un certo intervallo di tempo la fase di Testing. Una settimana dopo al soggetto veniva inviata una mail con il link diretto all'ultima fase dell'esperimento, quella del Test old-new finale. Una volta concluso il lavoro, è stata fornita una spiegazione del compito svolto ai partecipanti che lo hanno richiesto ed è stata data una risposta alle domande poste dai soggetti.

2.2.5 Questionari

Nel FSS III vengono presentate 24 situazioni spiacevoli e il soggetto deve valutare quanto ne verrebbe turbato. Il questionario valuta la sensibilità a situazioni a contenuto emotivo.

Il VVQ, Verbal-Visualizer questionnaire (Richardson, 1977) è costituito da 15 item a scelta forzata (vero o falso) in cui viene richiesto al soggetto la sua opinione rispetto alla frequenza, facilità o preferenza di comportamenti che prevedono l'uso di parole o di immaginazione.

Il QSSV, Questionario di strategie visive-verbali (Antonietti e Giorgetti, 1993), attraverso 18 item, con risposta di tipo Lickert a 5 punti, permette di ottenere una misura della tendenza ad impiegare nei processi cognitivi strategie o rappresentazioni di tipo visivo o verbale.

2.2.6 Registrazione dei dati comportamentali

Data la registrazione a distanza, si sono raccolti i soli dati comportamentali che sono stati analizzati mediante PST E-Prime (Schneider, Eschman, e Zuccolotto, 2002), registrando l'accuratezza. I dati comportamentali sono stati rilevati per entrambe le sessioni sperimentali; si è eseguito un confronto con l'esperimento in laboratorio concentrandosi sul riconoscimento old-new finale e un confronto tra la fase testing e la fase di riconoscimento finale del presente studio sui 14 soggetti.

2.2.7 Disegno sperimentale

Per valutare il ruolo del contenuto emozionale e del livello di nitidezza dell'immagine nella performance comportamentale abbiamo usato un disegno *within-subject*. Le analisi statistiche sono state effettuate per la seconda sessione sperimentale, quella avvenuta a distanza di una settimana. È stata adottata un'ANOVA univariata 3×2 , in quanto la nostra variabile blocco possiede, oltre al consueto integro e degradato, anche una quota di immagini che viene presentata solo una volta in fase di codifica, ossia che non prende parte al Testing (le no testing per l'appunto).

2.2.8 Risultati

Per quanto concerne la parte dedicata ai risultati del presente studio eseguito online, essendo esso un'implementazione del paradigma svolto precedentemente in laboratorio, ci occuperemo del confronto tra i due setting sperimentali, andando ad esplorare analogie e differenze dei dati comportamentali raccolti.

A distanza di una settimana dalla fase di codifica e di test, i partecipanti hanno eseguito un compito di riconoscimento *old-new*, in cui tutti gli stimoli sperimentali sono stati presentati integri (nella loro risoluzione originale). In particolare, in questa fase

- un terzo delle immagini è stato presentato per la prima volta (mai presentate prima)
- un altro terzo delle immagini era stato presentato durante la fase di Codifica (solo una volta sottoforma di gif)
- un terzo erano state prima presentate durante la fase di Codifica e poi riviste durante la fase di Testing e proprio in questa fase di Testing, una porzione era somministrata in modo integro e un'altra in modo degradato.

2.2.8.1 Risultati comportamentali nel Test di riconoscimento *Old-New*

Si è eseguita un'analisi sull'accuratezza per il campione di 14 partecipanti, tuttavia, non essendo l'accuratezza la misura ideale in un compito in cui vi è una scelta old/new perché non rivela la possibilità che i soggetti commettano un *bias* conservativo o non conservativo, si è ricorsi all'analisi sul d' (d' prime), che permette di rimuovere il rumore dal segnale registrato e ottenere una misura dipendente al netto dei bias dei soggetti (se ce ne sono).

La sensibilità del soggetto, indicizzata dal d' , è stata calcolata sottraendo gli Hit standardizzati dai Falsi Allarmi standardizzati per ciascuna condizione, come stato fatto per il test intermedio. La formula è: $[d' = z(H) - z(F)]$. Questa formula serve per scorporare la quota di rumore da quella di segnale. In riferimento alla Teoria della detezione del segnale

(TDS), in un compito di riconoscimento *old-new*, come quello adottato in questo esperimento, i possibili esiti di risposta sono quattro (tab.3):

		RISPOSTA del soggetto	
		SI	NO
Segnale (condizione di stimolazione)	SN (segnale presente)	H (Hit)	M (Miss)
	N (segnale assente)	FA (false alarm)	CR (correct rejection)

Tab 3: possibili esiti di risposta in un compito di riconoscimento *Old-New*. SN corrisponde al segnale grezzo contenente una proporzione di rumore mentre N corrisponde al solo rumore (noise)

- **Hit:** quando uno stimolo viene riconosciuto come già visto ed effettivamente è stato già presentato;
- **Correct rejection:** quando il soggetto seleziona il tasto nuovo e lo stimolo in questione è effettivamente nuovo;

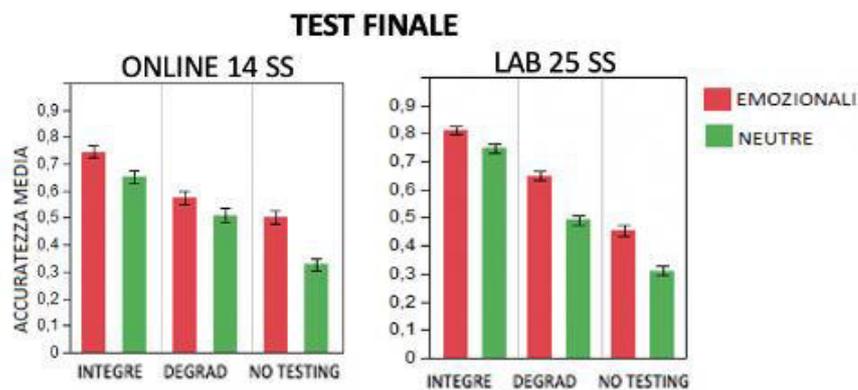
In entrambi i casi il soggetto ha risposto correttamente.

- **False alarms:** quando il soggetto classifica come vecchio uno stimolo che in realtà non è mai stato presentato prima;
- **Miss:** quando il soggetto non riconosce lo stimolo presentato come già visto;

In entrambi i casi il soggetto NON ha risposto correttamente.

Per quanto riguarda la fase finale di riconoscimento *old/new*, si è andati ad eseguire l'ANOVA 3x2 avente come fattori la condizione sperimentale (integra, degradata no testing) e la valenza (emotiva vs. neutra).

Da questa analisi sull'accuratezza, per il campione di 14 partecipanti, emerge un effetto principale della condizione sperimentale [F(2, 26) = 72.886 , p < .001, η_p^2 =.849] come per il campione da 25 soggetti dell'esperimento svolto in laboratorio [F(2, 48) =277.404 , p < .001, η_p^2 =.920].



A

Fig.A: si riferisce ai 14 partecipanti dell'esperimento online e i 25 dell'esperimento svolto in laboratorio. Viene riportata l'accuratezza con relativa barra d'errore in un test di riconoscimento *old-new*, una settimana dopo la fase di Codifica. Gli stimoli sono stati presentati tutti nella loro versione integra.

Nel grafico A viene evidenziato l'effetto principale della valenza [F(1, 13) = 24.267, p =.001, η_p^2 =.651] come anche per i 25 soggetti dell'esperimento in laboratorio [F(1, 24) = 36,432, p <.001, η_p^2 =.603] e anche l'interazione della condizione sperimentale con la valenza, per i 14 soggetti dell'esperimento online, risulta significativa

[F(2, 26) = 6.666, p =.005, η_p^2 =.339] come per il campione di 25 soggetti [F(2, 48) = 5.590, p =.007, η_p^2 =.189].

Inoltre, si è eseguita un'ANOVA 3x2 sul d', per i motivi descritti pocanzi, in cui ci sono sempre la condizione sperimentale (integra, degradata, no testing) e la valenza (emotiva vs. neutra) come fattori.

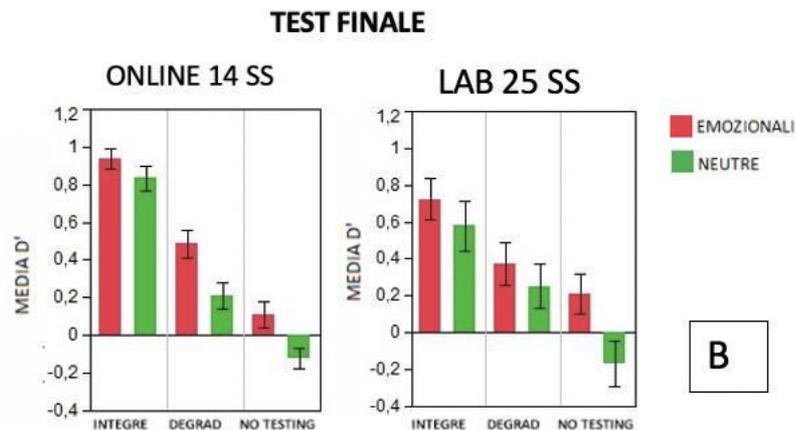


Fig. B: Grafico d' dei 14 partecipanti dell'esperimento online e dei 25 partecipanti, di quello svolto in laboratorio; entrambi riferiti al test finale

Si ha un effetto principale per la valenza [F(1,13) =10.614 , p =.006, η_p^2 =.449] come anche per i 25 soggetti dell'esperimento in laboratorio [F(1, 24) = 24.236, p <.001, η_p^2 =.502]. Anche qui, l'andamento è replicato anche se statisticamente meno forte (p=.006).

Inoltre, l'effetto principale della condizione risulta significativo [F(2, 26) = 63.283 , p <.001, η_p^2 =.830] come per i 25 soggetti [F(2, 48) =276.581 , p <.001, η_p^2 =.920]. L'interazione, infine, tra condizione e valenza, risulta anch'essa significativa [F(2,26) =6.811, p=.004, η_p^2 =.344]

come per i 25 soggetti del precedente esperimento [F(2, 48) = 5.589, p = .007, η_p^2 = .189].

La significatività nell'interazione dei 14 soggetti consente di far dei confronti e anche qui, all'interno delle varie condizioni sperimentali per le due valenze, appaiono essere tutte significative (p < .001) tranne per le degradate vs le no testing emozionali [F(1,13)=6.250, p < .027, η_p^2 = .325].

Per i 14 soggetti, andando a confrontare le due valenze all'interno dei diversi livelli di degradazione, emozionali e neutre non sono significativamente diverse nella condizione integre [F(1,13)=2.909, p = .112, η_p^2 = .183] e nemmeno nelle degradate [F(1,13)=2.008, p = .180, η_p^2 = .134]. Mentre, risulta significativa solo la differenza tra le emozionali e le neutre nelle No testing [F(1,13)=21.098, p = .001, η_p^2 = .619] come per l'accuratezza.

Per i 25 soggetti invece, in tutte le condizioni sperimentali la differenza tra emozionali e neutre risulta essere significativa, anche se, per le integre alla soglia della significatività [F(1,24)=4.501, p = .044, η_p^2 = .158]

2.2.8.2 Risultati comportamentali nel Test Intermedio

Per quanto riguarda la fase di testing intermedio, si è andati ad eseguire l'ANOVA 2x2 sul d', avente come fattori la condizione sperimentale (integra, degradata) e la valenza (emotiva vs. neutra).

Si riscontra un effetto principale per la condizione sperimentale [F(1,13) = 10.969 , p = .006, η_p^2 = .458] e un effetto principale valenza [F(1,13) = 6.06, p = 0.29, η_p^2 = .318] in cui si nota il vantaggio per le immagini emozionali rispetto alle neutre specificatamente all'interno

delle immagini integre. L'interazione tra condizione e valenza risulta essere significativa [$F(1,13) = 8.753, p = .011, \eta_p^2 = .405$] e questo permette di far dei confronti.

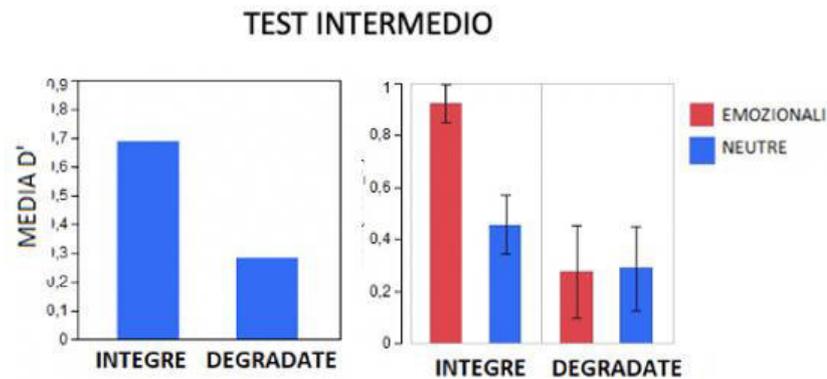


Fig.2.5: Il **grafico** è riferito alla fase testing dell'esperimento online per i 14 partecipanti

Confrontando le due condizioni sperimentali per le due valenze, risulta evidente il vantaggio dell'emozionalità nella versione integra degli item [$F(1, 13) = 15,361, p = .002, \eta_p^2 = .542$] rispetto alla controparte neutra e, una significativa differenza tra integre e degradate tenendo fissa l'emozionalità [$F(1, 13) = 17.746, p = .001, \eta_p^2 = .577$]. Non si nota una differenza significativa tra emozionali e neutre all'interno della condizione sperimentale degradata [$F(1, 13) = .008, p = .930, \eta_p^2 = .001$] e neanche una grande differenza tra le neutre integre e quelle degradate [$F(1, 13) = 1.455, p = .249, \eta_p^2 = .101$].

Il calcolo del d' mostra un bias conservativo: il soggetto commette una quota elevata di falsi allarmi. Lo stimolo degradato appare quindi già svantaggiato in fase di testing.

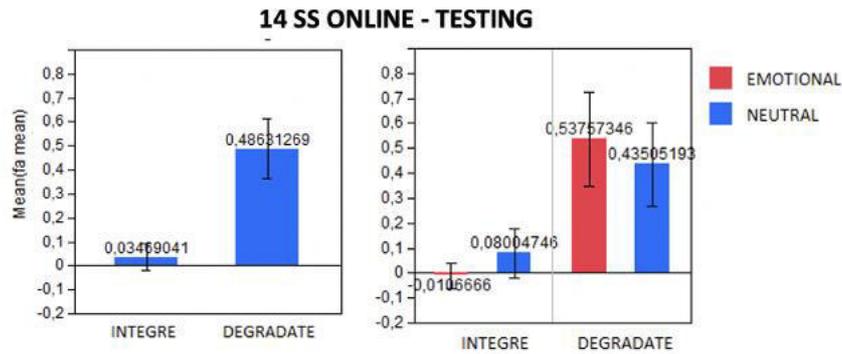
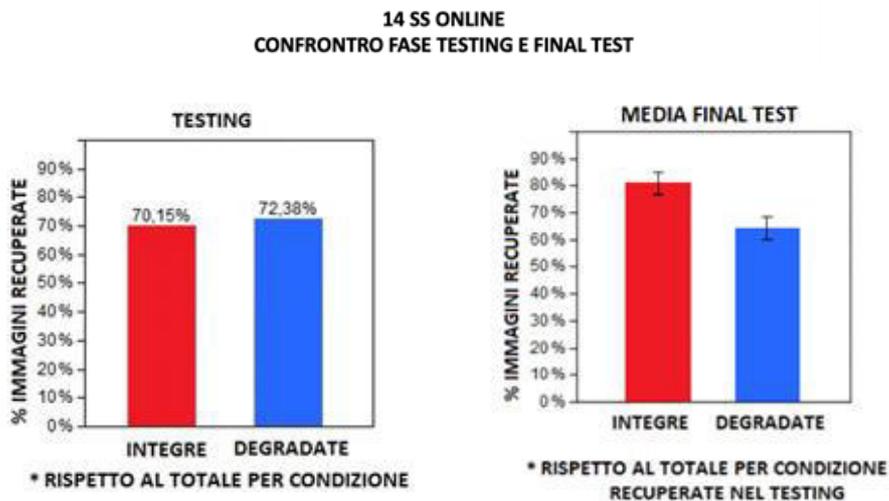


Fig.2.6: falsi allarmi nel testing intermedio: i soggetti tendono a rispondere old nelle degradate

2.2.8.3 Confronto risultati Test intermedio vs Test di riconoscimento Old New

Si esegue un confronto prendendo entrambe le condizioni sperimentali e andando a considerarne il 100%, si è andati a guardare la percentuale di immagini recuperate correttamente nel testing che viene poi recuperata allo stesso modo nel Test di riconoscimento old-new finale.



C

Fig. C: percentuale delle immagini riconosciute correttamente sia nel test intermedio che in quello finale

Nel test intermedio, considerando le sole immagini OLD e prendendo singolarmente le due condizioni (integre e degradate) considerate come il 100% rispetto a cui valutare la sensibilità del soggetto nel rispondere, si è visto che entrambe presentano una percentuale di immagini riconosciute accuratamente intorno al 70% (integre 70,15% e degradate 72,38%).

Nel Final Test, considerando la quota di immagini riconosciute nel testing per condizione (integre/ degradate) ciascuna come totale (100%) e calcolando quante immagini sono state riconosciute anche nel test finale: delle integre sopravvivono 81,15% mentre le degradate subiscono una maggiore perdita, con 65,65% recuperate.

Confrontando la media delle due condizioni (integro e degradato), emerge un effetto principale [$F(1, 13) = 29,568$, $p < .001$, $\eta_p^2 = .695$].

3. DISCUSSIONE

Questa particolare versione online di un esperimento sul testing Effect nel dominio della memoria emozionale per scene naturali, già svolto nel laboratorio di Psicofisiologia di Parma, è stata creata per implementare il suddetto lavoro arrivando ad una versione utilizzabile da remoto e con l'obiettivo di ampliare la rosa di dati raccolti. Come già detto in precedenza, lo studio è stato sottoposto a 66 persone ma dalle analisi sono stati esclusi la maggior parte dei soggetti per problemi legati al caricamento durante le prove svolte dal proprio laptop o per aver interposto una distanza maggiore di 1 settimana tra la prima prova e la seconda (il test di riconoscimento old-new finale). Questo è sicuramente un grande svantaggio legato agli esperimenti online di una durata elevata

come lo studio qui presentato. Gli strumenti a disposizione in uno studio da remoto sono fondamentali per la sua riuscita e spesso, l'inadeguatezza degli stessi può portare a problematiche come questa. Inoltre, è da ricordare che la partecipazione di un soggetto ad un esperimento deve sempre essere sostenuta da una grande motivazione, e questo vale ancor di più per gli esperimenti da remoto in cui lo sperimentatore, una volta date le istruzioni, può far ben poco.

L'esperimento affrontato nell'elaborato è stato costruito e sviluppato in tre fasi: codifica, test intermedio e test finale fatto eseguire a distanza di una settimana. Le due principali manipolazioni nei confronti del materiale sono state la variazione del contenuto emozionale e quindi l'aspetto semantico dello stesso (il significato dell'immagine), e l'aspetto percettivo dell'immagine: il grado di nitidezza del materiale presentato. Il fenomeno del *testing effect*, infatti, si verifica attraverso la somministrazione di indizi (i.e. cue) che facilitano la rievocazione del materiale (condizione testing), in alternativa al materiale integro (condizione studio). Cercando di calcare le orme di un classico esperimento sull'effetto testing nel dominio verbale, nel suddetto esperimento con le immagini, nella fase Testing, si è utilizzato un compito di riconoscimento. Una parte del materiale è stato presentato come integro, mentre un'altra parte come degradato, fornendo così un *cue* impoverito in termini percettivi in grado di riattivare la traccia mnestica, senza dover ripresentare il materiale per intero ricordando che, nella fase di codifica iniziale, sono state presentate scene naturali dinamiche, le *gif*, in cui si passava da una traccia impoverita e degradata, alla stessa immagine completamente nitida, per poi tornare nuovamente alla versione degradata. Questo per creare una sorta di associazione tra un indizio degradato, un cue, e il target nitido.

Nel test di riconoscimento finale, sia per l'esperimento online che per quello in laboratorio, si riscontra un vantaggio per le immagini integre rispetto alle degradate, vantaggio visto sia nell'analisi dell'accuratezza che nel d' (le due analisi tendono ad avere andamenti simili), rispondendo così alla nostra prima ipotesi, ovvero che una maggior ricchezza di dettagli porti ad un vantaggio in una fase di test finale. Ci si è chiesti se questo vantaggio per le immagini integre ci fosse già nel testing e, guardando l'accuratezza delle sole OLD nel testing, sembrerebbe non ci sia un vantaggio per le sole integre, ma che i soggetti abbiano uno stesso grado di accuratezza per entrambe le condizioni di nitidezza. Tuttavia, guardando al d' e ai FA del testing, i soggetti hanno già in un intervallo di tempo breve (quello che trascorre tra la codifica e il testing), difficoltà a riconoscere le immagini degradate (per le degradate, tendono a rispondere old anche se sono new). Pertanto, l'accuratezza di questa quota di immagini risulta essere la stessa delle integre, ma la sensibilità è diversa, sono maggiori i falsi allarmi. Quindi, già in fase testing si riscontra un vantaggio per le immagini integre, quelle ricche di dettagli. Nella fase di Testing intermedio, per l'appunto, i partecipanti hanno preso parte ad un test di riconoscimento in cui veniva presentato un blocco di stimoli integri ed uno di stimoli degradati, chiedendo loro di eseguire un test di riconoscimento vecchio-nuovo. Un compito di riconoscimento vecchio-nuovo con sole immagini integre viene fatto eseguire poi in fase finale.

Andare a far un confronto tra il test intermedio e quello finale, ha permesso quindi di rispondere ad uno dei quesiti posti all'inizio di questo esperimento ovvero: che tipo di contributo ha il testing sul test finale? Sappiamo che il testing è più efficace quanto più è alto il tasso di recupero (retrieval rate) nel testing stesso. I dati mostrano che, andando ad eseguire un confronto tra le due condizioni, prendendo le sole immagini Old, e tra

queste solo le immagini riconosciute accuratamente in questa fase dell'esperimento, il tasso di recupero tra immagini integre e degradate nel testing è pressoché lo stesso. Come è possibile vedere nel grafico C, la quota di immagini accurate si aggira intorno al 70% per entrambe le condizioni. Nel test finale risulta che in media circa l'81% delle le immagini integre vengono recuperate sulla totalità delle old accurate nel testing, e in media circa il 64% delle immagini degradate facendo emergere così, un vantaggio maggiore per le immagini integre ricordate accuratamente tra i due test e una percentuale minore di degradate.

Questa diversa sensibilità nella discriminazione a favore delle immagini integre sottolinea una differenza qualitativa: il recupero delle degradate è superficiale e questo non permette di consolidarne una traccia forte quanto quella delle integre. C'è un processo diverso coinvolto nel recupero delle degradate rispetto alle integre, in un caso più accurato (per le integre), nell'altro meno (per le degradate). Lo stimolo degradato appare quindi già svantaggiato in fase di testing. Le immagini degradate non raggiungono livelli di performance positiva al pari delle immagini integre. Potrebbe essere quindi che la natura del cue favorisca da sola un maggior o minor successo. Analizzando quelle ricordate con successo, sia integre che degradate, a parità di immagini ricordate nel testing, nel final test ne sopravvivono una quota maggiore di integre. La sensibilità di discriminazione nel Testing è diversa per le due condizioni: è più sensibile per le integre e ne favorisce una traccia più solida; mentre per le degradate, questo discrimination index è minore e influenza il test finale. Perciò il contributo del testing si riscontra in fase finale, nel test di riconoscimento old/new, sottolineando l'importanza della pratica del recupero; tuttavia, il consolidamento della traccia mnestica del materiale cambia a seconda del materiale stesso (della qualità del materiale). Il cue degradato ha quindi

successo nel recupero ma questo non è così efficace da far sì che ci sia un ri-consolidamento reale della traccia.

Nel test di riconoscimento finale, l'accuratezza di questa quota di immagini degradate risulta comunque migliore rispetto alle immagini presentate solo in codifica e non testate (soprattutto nella loro versione neutra, le *no testing*). Testare le immagini aiuta nel recupero esplicito (fase intermedia) e a combattere l'oblio e porta a prestazioni migliori sia per gli item integri che, anche se in misura minore, per quelli degradati (nel test finale). Allora è possibile supporre che la quantità di dettaglio vada a contrastare l'oblio. Infatti, alcuni dettagli specifici di un'immagine possono essere usati come *cue* in un test di riconoscimento (Nelson & Loftus, 1980) e la mancanza di questi dettagli nella versione degradata, porterebbe ad un peggioramento della performance.

Secondo il modello del "doppio processo", dal punto di vista qualitativo, infatti, i processi cognitivi che intervengono nel riconoscimento sono due: la *familiarity* e la *recollection*. Quando si viene esposti ad un certo stimolo, il riconoscimento dello stesso può essere basato sul ricordo vivido di dettagli e ricco di particolari contestuali (*recollection*), oppure su una sensazione di aver già fatto esperienza di quel determinato stimolo, la cosiddetta familiarità (*familiarity*) di quell'item (Yonelinas, 2002). Vi sono numerose prove a sostegno dell'esistenza di questi due processi di riconoscimento mnestico e in questo elaborato si può parlare di Retrieval Relay Race efficace, poiché già nel testing intermedio si nota una differenza nella sensibilità di riconoscimento delle immagini tra le integre e le degradate, forse attribuibile proprio ad un diverso tipo di elaborazione che si ha dello stimolo presentato.

Nel test finale di riconoscimento old-new, come si è detto in precedenza, vengono mostrate ai partecipanti tutte le immagini presentate in codifica, più altri stimoli mai somministrati prima. In questa fase le scene naturali sono state presentate nel loro formato originale ed è stata analizzata la proporzione di accuratezza in funzione della condizione: emozionalità e nitidezza e, questa manipolazione ha richiamato sicuramente in causa le *Teorie dello sforzo del recupero*. Secondo tali teorie al crescere della difficoltà del test intermedio, quello presentato dopo la codifica, cresca anche il beneficio che si riscontra nel test finale (Jacoby, 1978; Karpicke & Roediger, 2007). Pertanto, poiché un test di riconoscimento con stimoli degradati, risulta più complesso di uno con stimoli integri, ci si sarebbe aspettato che in fase finale, questo avesse prodotto un maggior beneficio, portando ad una miglior prestazione per quegli stimoli “incompleti”, ma non è stato così. Si è infatti verificato lo scenario opposto come visto già nell’esperimento in laboratorio. Non parliamo di un decadimento significativo della performance ma sicuramente non troviamo un reale beneficio per questi item privi di dettagli. Questo forse perché le immagini degradate in memoria non corrispondono ad un sottoinsieme delle immagini nitide (Wolfe, 2011). Nonostante ciò, presentare un *cue* impoverito in termini percettivi agisce sulla traccia di memoria, producendo comunque un tasso di riconoscimento maggiore rispetto a quanto emerge con la quota di immagini presentata solo una volta in codifica e non testate. Tuttavia, sono le immagini integre, infatti, quelle che godono di un vantaggio nel consolidamento della traccia mnemonica maggiore rispetto alle immagini degradate tra la fase di test intermedio e quella di test finale. È come se il consolidamento di queste immagini sia più profondo rispetto al ricordo “superficiale” dell’immagine presentata come degradata. La sensibilità

della discriminazione per le immagini degradate risulta essere inferiore rispetto alle integre; il cue degradato viene ricordato con successo ma non come accade per l'integro. Oltre a ciò, in condizioni ottimali la memoria di riconoscimento per immagini risulta davvero sorprendente (Nickerson, 1968; Standing, Conezio e Haber, 1970; Standing, 1973), dando luogo ad un fenomeno noto in letteratura come *effetto di superiorità pittorica*. I risultati ottenuti suggeriscono che, anche il formato delle immagini scelto in codifica utilizzato in entrambi gli esperimenti, ossia la presentazione di un'immagine sfocata che diventa nitida per poi sfocarsi nuovamente, possa andare a generare un decadimento più o meno profondo; presentare un'immagine integra in un formato dinamico tra due versioni degradate, può produrre infatti effetti di interferenza proattiva e retroattiva, ovvero dei fenomeni in cui l'elaborazione cognitiva non è accurata a causa di processi competitivi tra informazioni correnti e precedenti (McLeod, Saul, 2018).

Per quanto concerne il contenuto emozionale ci si è chiesti se, anche in questa versione online dell'esperimento, emergesse il vantaggio per la componente emozionale come nella maggior parte degli esperimenti offline. È emerso come le immagini a contenuto piacevole e quelle a contenuto spiacevole, vengano ricordate meglio rispetto a quelle dal carattere neutro, come gli interni di una casa o gli spazi all'aperto e coerentemente con la letteratura (Anderson et al., 2006; Ochsner, 2000; Ritchey et al., 2008; Sterpenich et al., 2009; Weymar et al., 2009), questi stimoli a contenuto emozionale resistono maggiormente all'oblio perché contengono appunto, elementi più salienti. Dai dati comportamentali si evince come il vantaggio per le immagini emozionali rispetto alle immagini a contenuto neutro, cresca al crescere della difficoltà del compito. Sebbene però la differenza tra emotive e neutre sia sempre

significativa, come visto nell'esperimento in laboratorio, guardando ai risultati del test intermedio e a quelli del test finale in questo esperimento, si può parlare sicuramente di un vantaggio per le integre emozionali rispetto alle neutre e alle degradate nel testing, ma questo effetto si attenua nel test finale. Qui troviamo un recupero maggiore delle neutre nella versione integra. Infatti, nell'esperimento online, a differenza dell'esperimento svolto in laboratorio, emozionali e neutre non appaiono così diverse nelle condizioni integre e degradate, differenza che si riscontra solo nelle No testing. Tenendo ferma la valenza, le immagini emozionali hanno un maggior impatto se viste solo una volta (no testing) rispetto all'essere riviste ma in modo degradato ($p < .027$ effetto piccolo). Sembra che il contenuto emozionale favorisca la memorabilità delle immagini soprattutto se viste solo una volta rispetto a quanto non accada per quelle riviste, sia in modo degradato che integro. Nonostante ciò, l'effetto dell'emozionalità nelle integre rimane significativo, confermando quanto visto nell'esperimento in laboratorio.

I dati suggeriscono che la valenza del materiale abbia un effetto significativo sulla capacità di memoria: questo influenza la memorabilità della traccia rispetto al materiale neutro. L'effetto risulta evidente soprattutto con gli item nella condizione *no testing* (e quindi apparsi solo in codifica e poi direttamente nel test finale) in cui c'è un vantaggio ancora maggiore a favore delle immagini a contenuto emozionale rispetto alle neutre. La valenza del materiale, soprattutto se di mezzo non c'è una fase testing, aiuta nella ritenzione e nel recupero della memoria.

In sintesi, i dati suggeriscono che l'effetto del test può essere generalizzato al dominio delle memorie di scene complesse, con una migliore discriminazione per gli elementi testati rispetto a quelli non testati. Il testing appare essere quindi un metodo efficace per aiutare nel

recupero e nel consolidamento mnestico, ma nel dominio delle immagini, presentare immagini degradate, cues incompleti, ne riduce l'efficacia.

L'esperimento svolto, in linea generale, va a confermare quanto si è visto nell'esperimento di laboratorio e mostra come l'utilizzo delle nuove tecnologie, nonostante molti scetticismi a riguardo, permetta di replicare fedelmente, seppur con alcune modifiche, paradigmi sperimentali svolti in laboratorio e consenta di riprodurli su piattaforme online in modo da poter aumentare la numerosità campionaria e di conseguenza anche l'affidabilità dei dati raccolti (rapporto segnale-rumore). Nell'elaborato si è visto così quanto il paradigma del Testing possa essere informativo nell'evidenziare i processi che sono alla base della memoria visiva per scene complesse. Pertanto, gli esperimenti futuri potrebbero esplorare come diverse manipolazioni percettive e contestuali delle immagini, o attraverso la manipolazione della modalità di svolgimento del compito richiesto durante la fase test (es. riconoscimento vs. test di richiamo, o memoria/categorizzazione), possa influire sul consolidamento e sul recupero della traccia della memoria visiva.

APPENDICE

A- Questionari

B- Script di MATLAB

C- Virtual Chinrest

D- Totale soggetti

A- Questionari

1. Consenso informato



LABORATORIO DI PSICOFISIOLOGIA

Prof. ssa Vera Ferrari
Plesso Biotecnologico Integrato
Via Volturmo, 39 – 43126 PARMA
www.unipr.it

Consenso Informato

Progetti di ricerca psicologica: LA MEMORIA FOTOGRAFICA E LE EMOZIONI

Gentile Signora/Signore

Prima di decidere liberamente se vuole partecipare a questo studio, LEGGA ATTENTAMENTE questo consenso informato e ponga al responsabile della ricerca tutte le domande che riterrà opportune al fine di essere pienamente informato degli scopi, delle modalità di esecuzione dell'esperimento e dei possibili inconvenienti connessi. La preghiamo di ricordare che questo è un progetto di ricerca e che la sua partecipazione è completamente volontaria. Lei si potrà ritirare in qualunque momento. Ogni informazione che otterremo da lei, inclusa la sua identità, sarà mantenuta riservata. Le ricordiamo che in caso avesse bisogno di delucidazioni su qualunque aspetto della procedura sperimentale, il responsabile della ricerca è la Prof.ssa Vera Ferrari, vera.ferrari@unipr.it, che insieme ai suoi collaboratori, sarà a Sua completa disposizione.

La ricerca ha lo scopo di studiare i meccanismi che sottendono il consolidamento e il recupero di informazioni presenti nella memoria a lungo termine, utilizzando materiale visivo con il quale tutti noi abbiamo una elevata familiarità, ossia immagini di scene naturali. I contenuti raffigurati riguarderanno scene di vita quotidiana, così come scene ad alto contenuto emotivo, quali scene erotiche, animali feroci e incidenti. In considerazione delle particolari immagini proposte, vi invitiamo alla massima cautela, e raccomandiamo di non esporre altre persone al materiale sperimentale, se non espressamente invitati a far parte della ricerca.

La partecipazione al progetto di ricerca comporta la compilazione di una scheda informativa (Anamnesi), tre brevi questionari, e il calcolo della distanza dallo schermo da mantenere per l'intera sessione sperimentale. L'esperimento prevede due sessioni che si terranno a distanza di una settimana l'una dall'altra. Lo svolgimento dell'esperimento al PC richiederà circa 60 minuti per sessione.

La partecipazione allo studio è volontaria e gratuita. I partecipanti sono liberi di ritirarsi dalla collaborazione in qualsiasi momento senza preavviso o motivazione specifica. I partecipanti hanno diritto a richiedere informazioni sui risultati e sull'esito della ricerca. L'elaborazione dei dati raccolti sarà condotta in modo da eliminare qualsiasi riferimento che possa permettere di ricollegare singole affermazioni a una determinata persona. I risultati della ricerca saranno pubblicati in forma riassuntiva e in nessun caso eventuali brevi citazioni saranno riconducibili a singole persone.

Per qualsiasi informazione e chiarimento su questo studio o per qualsiasi necessità può rivolgersi alla Prof.ssa Vera Ferrari (vera.ferrari@unipr.it) che è a sua disposizione per ulteriori informazioni o chiarimenti.

Consenso informato alla partecipazione allo studio

Il/la sottoscritto/a _____ residente in _____, via _____

DICHIARA

di aver letto il suddetto foglio informativo ricevuto, di aver compreso sia le informazioni in esso contenute sia le informazioni fornite in forma orale dal personale addetto al progetto di ricerca "Testing Effect

Tel. Laboratorio ++ 39 0521 034946
E-mail : vera.ferrari@unipr.it



**UNIVERSITÀ
DI PARMA**

DIPARTIMENTO DI MEDICINA E CHIRURGIA

LABORATORIO DI PSICOFISIOLOGIA

Prof. ssa Vera Ferrari
Plesso Biotecnologico Integrato
Via Volturmo, 39 – 43126 PARMA
www.unipr.it

nel dominio della memoria per immagini” e di aver avuto ampio tempo ed opportunità di porre domande ed ottenere risposte soddisfacenti dal personale addetto;

- di aver compreso che la partecipazione allo studio è del tutto volontaria e libera, che ci si potrà ritirare dallo studio in qualsiasi momento, senza dover dare spiegazioni e senza che ciò comporti alcuno svantaggio o pregiudizio;
- di aver compreso la natura e le attività che la partecipazione allo studio comportano e i relativi rischi;
- di aver compreso che la partecipazione a questo studio non comporterà il riconoscimento di alcun vantaggio di natura economica diretto o indiretto.

Conseguentemente, il/la sottoscritto/a

ACCONSENTE NON ACCONSENTE

a partecipare allo studio, nella consapevolezza che tale consenso è manifestato liberamente ed è revocabile in ogni momento senza che ciò comporti alcuno svantaggio o pregiudizio.

(luogo e data)

(firma)

firma del responsabile della ricerca:
Prof.ssa Vera Ferrari

Tel. Laboratorio ++ 39 0521 034946
E-mail : vera.ferrari@unipr.it

QUESTIONARIO FSS

ISTRUZIONI: Le voci di questo questionario si riferiscono a cose o esperienze che possono causare paura o altri sentimenti spiacevoli. Metta una X nel quadrato che meglio descrive quanto ne sei turbato/a

Nella versione cartacea troviamo una griglia con le varie voci a cui corrispondono delle risposte tra cui può scegliere il partecipante all'esperimento:

		Per nulla	Poco	Abbastanza	Molto	Moltissimo
1.	Ferite aperte	<input type="checkbox"/>				
2.	Gente morta	<input type="checkbox"/>				
3.	Vedere una persona che fa la prepotente con gli altri	<input type="checkbox"/>				
4.	Farsi fare delle iniezioni	<input type="checkbox"/>				
5.	Cimici	<input type="checkbox"/>				
6.	Insetti volanti	<input type="checkbox"/>				
7.	Vedere altre persone che ricevono iniezioni	<input type="checkbox"/>				
8.	Armi	<input type="checkbox"/>				
9.	Sporcizia	<input type="checkbox"/>				
10.	Insetti che strisciano	<input type="checkbox"/>				
11.	Sangue	<input type="checkbox"/>				
12.	Ampi spazi aperti	<input type="checkbox"/>				
13.	Malati	<input type="checkbox"/>				
14.	Assistere ad un intervento chirurgico	<input type="checkbox"/>				
15.	Luoghi chiusi	<input type="checkbox"/>				
16.	Serpenti	<input type="checkbox"/>				
17.	Extrasistole (perdita di un battito cardiaco)	<input type="checkbox"/>				
18.	Topi	<input type="checkbox"/>				
19.	Ragni	<input type="checkbox"/>				
20.	Pipistrelli	<input type="checkbox"/>				

Nel caso del suo corrispettivo online troviamo schermate come questa:

Questionario FSS
Quello che ti verrà
presentato è un
questionario che si riferisce
a cose o esperienze che
possono causare paura o
altri sentimenti spiacevoli.
premi qui per procedere

Rispondi digitando un numero
da 1 a 5 secondo l'intensità che
meglio descrive quanto ne sei
turbato/a

1: per nulla
2: poco
3: abbastanza
4: molto
5: moltissimo

premi qui per procedere

Premi la barra
spaziatrice quando sei
pronto ad iniziare

Insetti che strisciano

Sporcizia

Cimici

Gente morta

Armi

Vedere altre persone
che ricevono iniezioni

“VERBALIZER-VISUALIZER QUESTIONNAIRE”

(VVQ) di A. Richardson

«Legga con attenzione ciascuna delle affermazioni sotto riportate.

Per ogni affermazione prema il tasto che corrisponde a ‘vero’ o ‘falso’ facendo riferimento alle Sue preferenze, alle Sue opinioni e a quanto da Lei abitualmente compiuto».

Nella versione online, il questionario si presenta in questo modo (es. di alcune domande)

The image displays a sequence of five grey rectangular screens from an online questionnaire. The first screen shows the title 'VERBALIZER-VISUALIZER QUESTIONNAIRE (VVQ) di A. Richardson' and instructions to read statements carefully. The second screen explains the response format: 'V' for 'VERO' and 'F' for 'FALSO'. The third screen instructs to press the spacebar to start. The fourth and fifth screens show two sample questions: 'Trovo difficile, chiudendo gli occhi, costruirmi un'immagine mentale del volto di un amico' and 'I miei sogni ad occhi aperti sono talvolta così vividi da avere la sensazione di vivere realmente la scena che sto immaginando.'

VERBALIZER-VISUALIZER
QUESTIONNAIRE (VVQ)
di A. Richardson

Leggi con attenzione ciascuna delle
affermazioni riportate

premi qui per procedere

Per ogni affermazione, facendo
riferimento alle tue preferenze,
alle tue opinioni e a quanto da
te abitualmente compiuto, premi
'v' per la risposta 'VERO' o 'f'
per la risposta 'FALSO'

premi qui per procedere

Premi la barra
spaziatrice quando sei
pronto ad iniziare

Trovo difficile,
chiudendo gli occhi,
costruirmi
un'immagine mentale
del volto di un amico

Le mie capacità di
immaginazione visiva
sono superiori alla
media

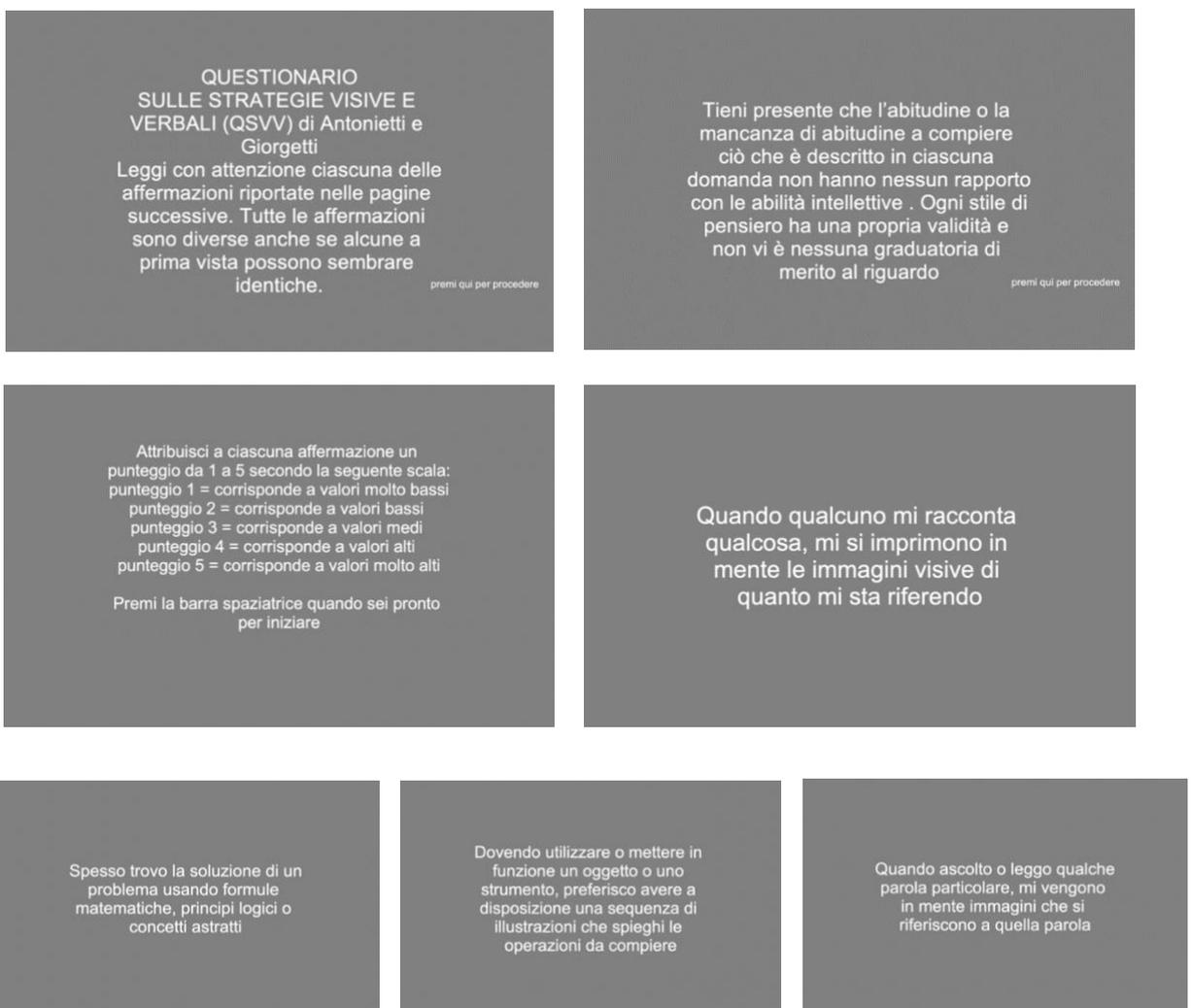
I miei sogni ad occhi
aperti sono talvolta
così vividi da avere la
sensazione di vivere
realmente la scena
che sto immaginando.

“QUESTIONARIO SULLE STRATEGIE VISIVE E VERBALI” (QSVV) di Antonietti e Giorgetti

«Legga con attenzione ciascuna delle affermazioni riportate nelle pagine successive. Tutte le affermazioni sono diverse anche se alcune a prima vista possono sembrare identiche.

Per ogni affermazione stabilisca in che misura è Sua abitudine compiere quanto da essa descritto. A tal fine, attribuisca a ciascuna affermazione un punteggio da 1 a 5.

Nella versione online le varie schermate si presentano in questo modo:



B - Script di MATLAB

```
degrad = 4.5;
intact = 8.0;
nsteps = 20;
step = (intact-degrad)/(nsteps-1); levels = [degrad:step:intact];

ExpTime = 0.1; ExpTime1 = 2; ExpTime2 = 2; ExpTimeMid = 1;
MidPic =nsteps +1;

%%%%%%%%%% end config
```

La prima parte definisce la degradazione minima e massima, e gli step tra l'una e l'altra; in questo caso parte da 4.5, arriva a 8, e fa un totale di 20 immagini.

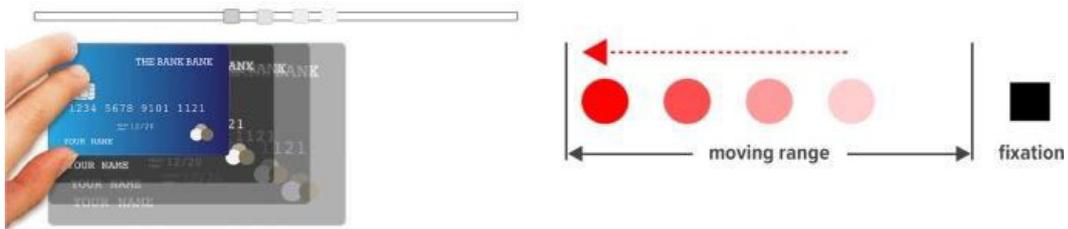
La seconda parte invece definisce il timing del video
exptime1 è la durata della prima immagine (la più degradata)

exptime2 è la durata dell'immagine finale (di nuovo la più degradata)
exptimemid è la durata dell'immagine a metà (quella intatta)
e midpic definisce qual è l'immagine intatta (il numero di step +1)
exptime e basta (senza mid, 1 o 2) è la durata di tutti gli altri frame
please cambia la riga MidPic = 21

```
MidPic =nsteps +1;
```

C- Virtual Chinrest

Serve a calcolare il tasso di conversione da pixel a cm per il display del computer di un partecipante. Viene chiesto ai partecipanti di posizionare una carta di credito o un altro oggetto che abbia quella stessa dimensione, sullo schermo del proprio computer e si chiede loro di ridimensionare un'immagine presentata, fino a quando quella non è della stessa dimensione della carta di credito o dell'oggetto dalla dimensione simile. Per misurare la distanza di visione del partecipante dal proprio schermo si usa un'attività di punti ciechi. Ai partecipanti viene chiesto di concentrarsi su un quadrato nero sullo schermo con l'occhio destro chiuso, mentre un punto rosso si muove ripetutamente da destra a sinistra. Viene chiesto loro di premere la barra spaziatrice ogni qual volta percepiscono che il punto rosso sia scomparso. Questa parte consente al plugin di utilizzare la distanza tra il quadrato nero e il punto rosso per stimare la distanza del partecipante dal monitor. Questa stima presuppone che il punto cieco si trovi a 13.5° .



D-Totale soggetti

Hanno partecipato all'esperimento 66 soggetti di cui solo 14 soggetti (3, 6, 12, 14, 19, 21, 33, 37, 47, 56, 57, 63, 64, 66) sono stati presi in considerazione per le analisi.

Motivi dell'esclusione dei soggetti (52 soggetti)

- S2 S16 bias nelle risposte (rispondono sempre uguale per più di 30 item);
- S4 S7 S8 S9 S13 S15 S28 S30 S31 S46 S54 S55 S62 non riescono a caricare l'esperimento, dall'inizio o da un certo punto in poi variabile (può capitare con esperimenti molto lunghi e il cui formato non viene sopportato dal proprio sistema operativo, come MacOS);
- S10 output contenente solo prima parte gif.;
- S11 output mancante della prima parte delle gif.;
- S17 S18 S23 S24 S25 S26 S27 S29 S32 S45 non eseguono seconda parte;
- S20 S22 S48 S53 S58 non attendono il salvataggio dei dati come richiesto nelle istruzioni;
- S5 S34 S35 S36 S38 S40 S49 S50 S52 S59 S60 S65 eseguono la seconda parte dopo la settimana prestabilita;
- S28 S39 S41 S42 S43 S44 S51 S61 non accedono al link inviato;

Bibliografia

Arnold K. M., McDermott K. B. (2013b). Test-potentiated learning: distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 940–945. doi:10.1037/a0029199.

Bartlett J. C., Tulving E. (1974). Effects of temporal and semantic encoding in immediate recall upon subsequent retrieval. *Journal of Verbal Learning & Verbal Behavior*, 13(3), 297–309. [https://doi.org/10.1016/S0022-5371\(74\)80066-6](https://doi.org/10.1016/S0022-5371(74)80066-6)

Bradley, Margaret M.; Greenwald, Mark K.; Petry, Margaret C.; Lang, Peter J. (1992). Remembering pictures: Pleasure and arousal in memory.. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2), 379–390. doi:10.1037/0278-7393.18.2.379

Buchanan T.; Smith J. L. (1999). Using the Internet for psychological research: Personality testing on the World Wide Web. , 90(1),125144. doi:10.1348/000712699161189

Buchanan T.(2000). Psychological Experiments on the Internet || Potential of the Internet for Personality Research. , (), 121–140. doi:10.1016/b978-012099980-4/50006-x

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1563–1569. doi:10.1037/a0017021.

Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, 34, 268–276. doi:10.3758/BF03193405.

Carrier M., Pashler H.(1992). The influence of retrieval on retention, 20(6), 633–642. doi:10.3758/bf03202713

Castelhano M., Fernandes S., Theriault J.,(2018). Examining the Hierarchical Nature of Scene Representations in Memory. *Journal of Experimental Psychology Learning Memory and Cognition* 45(9), DOI:10.1037/xlm0000660

- Dandurand F.; Shultz T. R.; Onishi K. H. (2008). Comparing online and lab methods in a problem-solving experiment. , 40(2), 428–434.doi:10.3758/brm.40.2.428
- Goldstein A. G.; Chance J. E. (1971). Visual recognition memory for complex configurations. , 9(2), 237–241.doi:10.3758/bf03212641
- Gosling S.D.; Vazire, S.; Srivastava, S.; Oliver J. (2004). Should We Trust Web-Based Studies? A Comparative Analysis of Six Preconceptions About Internet Questionnaires. *American Psychologist*, 59(2), 93–104.doi:10.1037/0003-066X.59.2.93
- Henderson L.; Devine K.; Weighall A.; Gaskell G. (2015). When the daffodil flew to the intergalactic zoo: Off-line consolidation is critical for word learning from stories. *Developmental Psychology*, 51(3), 406–417.doi:10.1037/a0038786
- Henderson, L.M.; James, E. (2018). Consolidating new words from repetitive versus multiple stories: Prior knowledge matters. *Journal of Experimental Child Psychology*, 166(), 465–484. doi:10.1016/j.jecp.2017.09.017
- Hockley William E. (2008). The picture superiority effect in associative recognition. , 36(7), 1351-1359.doi:10.3758/mc.36.7.1351
- Hofstein A.; Lunetta V. N. (1982). The Role of the Laboratory in Science Teaching: Neglected Aspects of Research. *Review of Educational Research*, 52(2), 201–217. doi:10.2307/1170311
- Horst, Jessica S. (2013). Context and repetition in word learning. *Frontiers in Psychology*, 4(), –.doi:10.3389/fpsyg.2013.00149
- Houston-Price C.; Howe J. A.; Lintern N. J. (2014). Once upon a time, there was a fabulous funambulist...: what children learn about the “high-level” vocabulary they encounter while listening to stories. *Frontiers in Psychology*, 5(), –.doi:10.3389/fpsyg.2014.00075

- James E.; Gaskell M.; Weighall A.; Henderson L. (2017). Consolidation of vocabulary during sleep: The rich get richer?. *Neuroscience & Biobehavioral Reviews*, 77(),1–13.doi:10.1016/j.neubiorev.2017.01.054
- James W. (1890). *The principles of psychology*, Vol I. || . , 10.1037/10538-000(), –.doi:10.1037/10538-000
- Evans K.K., Baddeley A.(2018) Intention, attention and long-term memory for visual scenes: It all depends on the scenes, <https://doi.org/10.1016/j.cognition.2018.06.022>
- Juvrud J.; Gredebäck G.; Åhs F.; Lerin N.; Nyström P.; Kastrati G.; Rosén J. (2018). The Immersive Virtual Reality Lab: Possibilities for Remote Experimental Manipulations of Autonomic Activity on a Large Scale. *Frontiers in Neuroscience*, 12(), 305–.doi:10.3389/fnins.2018.00305
- Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological Research Online: Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist*, 59(2), 105–117. <https://doi.org/10.1037/0003-066X.59.2.105>
- Lachman R.; Laughery K.R. (1968). Is a test trial a training trial in free recall learning? *Journal of Experimental Psychology*, 76(1, Pt.1), 40–50. doi:10.1037/h0025278
- Levie W. H; Hathaway S. N. (1988). Picture Recognition Memory: A Review of Research and Theory. *Journal of Visual Verbal Language*, 8(1), 6–45.doi:10.1080/23796529.1988.11674426
- Li Q.; Joo, S. J.; Yeatman J. D.; Reinecke K. (2020). Controlling for Participants' Viewing Distance in Large-Scale, Psychophysical Online Experiments Using a Virtual Chinrest. *Scientific Reports*, 10(1), 904–. doi:10.1038/s41598-019-57204-1
- Marslen-Wilson W. D; Welsh A. (1978). Processing interactions and lexical access during word recognition in continuous speech. , 10(1), 29–63.doi:10.1016/0010-0285(78)90018-x

McDermott K. B. (2021). Practicing Retrieval Facilitates Learning. *Annual Review of Psychology*, 72(1), annurev-psych-010419-051019–.doi:10.1146/annurev-psych-010419-051019

Nedic, Z.; Machotka, J.; Nafalski, A. (2003). [IEEE 33rd Annual Frontiers in Education, 2003. FIE 2003. - Westminster, Colorado, USA (Nov. 5-8, 2003)] 33rd Annual Frontiers in Education, 2003. FIE 2003. - Remote laboratories versus virtual and real laboratories., 1(), T3E_1–T3E_6.doi:10.1109/FIE.2003.1263343

Nickerson, R. S. (1968). A note on long-term recognition memory for pictorial material. 11, 58–58. doi:10.3758/bf03330991

Olson R.K.J., Banaji M., Bruckman A., Cohen J., Couper M.,(2004). Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet., 59(2):105-17. doi: 10.1037/0003-066X.59.2.105.

Peirce, Jonathan W (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2(), –.doi:10.3389/neuro.11.010.2008

Peng Z. E., Buss E., Shen Y., et al. (2020) Remote testing for psychological and physiological acoustics: Initial report of the P&P Task Force on Remote Testing. doi: 10.1121/2.0001409

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, 81(1), 10–15. <https://doi.org/10.1037/h0027470>

Rawson K. A.; Dunlosky J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough?. *Journal of Experimental Psychology: General*, 140(3), 283–302. doi:10.1037/a0023956

Reips U.-D. (2000). Psychological Experiments on the Internet || The Web Experiment Method(), 89–117.doi:10.1016/b978-012099980-4/50005-8

- Roediger, Henry L.; Karpicke, Jeffrey D. (2006). The Power of Testing Memory .. *Perspectives on Psychological Science*, 1(3), 181–210. doi:10.1111/j.1745-6916.2006.00012.x
- Roediger, H. L., & Karpicke, J. D. (2006a). Test-enhanced learning: taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255. doi:10.1111/j.1467-9280.2006.01693.x.
- Roediger, H. L., & Karpicke, J. D. (2006b). The power of testing memory: basic research and implications for educational practice. *Perspectives on Psychological Science*, 1, 181–210. doi:10.1111/j.1745-6916.2006.00012.x.
- Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011a). Test-enhanced learning in the classroom: long-term improvements from quizzing. *Journal of Experimental Psychol*
- Roediger, H. L., Butler A.C. (2011). The critical role of retrieval practice in long-term retention, 15(1), 0–27. doi:10.1016/j.tics.2010.09.003
- Rowland, Christopher A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect.. *Psychological Bulletin*, 140(6), 1432–1463. doi:10.1037/a0037559
- Standing, L., Conezio, J. & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. 19, 73–74. doi:10.3758/bf03337426
- Tamara van Gog; Sweller J. (2015) Not New, but Nearly Forgotten: the Testing Effect Decreases or even Disappears as the Complexity of Learning Materials Increases, 27:247–264 DOI 10.1007/s10648-015-9310-x
- Toppino, T. C. & Cohen, M. S. (2009). The testing effect and the retention interval: que- stions and answers. *Experimental psychology*, 56, 252–257. doi:10.1027/1618- 3169.56.4.252
- Tulving, E., & Arbuckle, T.Y. (1966). Input and output interference in short-term associative memory. *Journal of Experimental Psychology*, 72, 145–150.

Wolfe, J. M. & Kuzmova, Y. I. (2011). How many pixels make a memory? Picture memory for small pictures. 18, 469–475. doi:10.3758/s13423-011-0075-z

Yonelinas, A. P. (2002a). Components of episodic memory: the contribution of recollection and familiarity. doi:10.1093/acprof:oso/9780198508809.003.0003

Yonelinas, A. P. (2002b). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. 46, 441–517. doi:10.1006/jmla.2002.2864