**Applicazione del Modello di Kirkpatrick per la Valutazione delle Competenze Infermieristiche in Sala Operatoria: uno Studio Pilota Basato su Gamification Interattivi in HTML**

Autori: Caterina Cicala1, Manuela Cavalletti2, Carmen Nuzzo3, Stefano Pepe4,

1 Responsabile Formazione Infermieri Fondazione Policlinico Gemelli Largo F. Vito, 1 - 00135 Roma - Italia

2 Coordinatore di dipartimento Fondazione Policlinico Gemelli Largo F. Vito, 1 - 00135 Roma – Italia

2 Coordinatore U.O.C. SITRA Fondazione Policlinico Gemelli Largo F. Vito, 1 - 00135 Roma - Italia

2 Infermiere di Sala Operatoria Fondazione Policlinico Gemelli Largo F. Vito, 1 - 00135 Roma - Italia

**Abstract**

L’inserimento di infermieri neoassunti in sala operatoria è un momento delicato che richiede strumenti mirati per valutare rapidamente conoscenze e abilità. Questo studio pilota propone un modello formativo innovativo che unisce il modello di Kirkpatrick alla tecnica di Gamifification interattivo in HTML (Mindex, 20 domande), somministrato in modalità blended.

L’obiettivo è misurare l’efficacia della formazione nei quattro livelli di Kirkpatrick: soddisfazione, apprendimento, comportamento e impatto organizzativo. La raccolta e l’analisi dei dati avvengono in tempo reale tramite dashboard dinamiche, con rappresentazioni grafiche utili al monitoraggio individuale e di gruppo.

Il campione comprende 8 infermieri neoassunti presso un grande ospedale pubblico, al primo anno di attività clinica in blocco operatorio. La metodologia prevede selezione dei partecipanti, randomizzazione, somministrazione del quiz e confronto pre/post formazione.

I risultati preliminari mostrano un miglioramento significativo nei livelli 1-3: soddisfazione per l’approccio digitale, incremento misurabile delle conoscenze e maggiore consapevolezza nell’applicazione delle procedure. L’impatto organizzativo (livello 4) richiede ulteriori studi su campioni più ampi e tempi di osservazione più lunghi.

I dati raccolti evidenziano differenze significative tra due gruppi, con punteggi mediamente superiori e più omogenei nel Gruppo 2 rispetto al Gruppo 1, confermando l’efficacia del modello formativo. Lo studio supporta l’uso di quiz HTML come strumento versatile, scalabile e replicabile in altri ambiti sanitari, favorendo percorsi di crescita professionale e maggiore sicurezza per il paziente.

**Parole chiave:** Kirkpatrick, Infermieri, Sala Operatoria, Valutazione Competenze, HTML, Quiz Interattivo, Formazione, Gamification

## Introduzione

La valutazione dell’efficacia delle attività formative riveste un ruolo centrale in ambito sanitario, in particolare nella formazione degli infermieri di sala operatoria, dove precisione, sicurezza e tempestività sono requisiti imprescindibili. Tra i modelli di riferimento, quello proposto da **Kirkpatrick (1994)** si è affermato come framework di elezione grazie alla sua struttura articolata in quattro livelli: **Reazione**, **Apprendimento**, **Comportamento** e **Risultati** [2].

Recenti evidenze ne hanno confermato la validità in contesti clinici complessi: **Liu et al. (2025)** hanno applicato il modello in un programma standardizzato per nuovi infermieri, evidenziando miglioramenti sia nella soddisfazione (Livello 1), sia nelle competenze (Livello 2), fino a ricadute osservabili sulla pratica clinica (Livello 3) e sugli outcome organizzativi (Livello 4) [1]. Analogamente, **Li et al. (2020)** hanno dimostrato, durante la pandemia da COVID-19, un impatto positivo del modello in contesti di chirurgia d’urgenza, con significativi incrementi di gradimento e competenze pratiche [3].

La letteratura di revisione (de Miranda et al., 2024) segnala tuttavia una maggiore applicazione dei primi tre livelli, mentre il Livello 4 rimane spesso sottoutilizzato a causa di difficoltà metodologiche e necessità di tempi e risorse dedicate [2]. In parallelo, l’ambito dell’e-learning ha iniziato a integrare al modello tradizionale la valutazione del **ROI (Return on Investment)**: **Pandey (2021)** propone infatti di associare agli esiti formativi anche indicatori economici, così da giustificare gli investimenti in tecnologie educative e formazione digitale [4].

Nel contesto operatorio, l’adozione del modello di Kirkpatrick, arricchito da strumenti digitali interattivi e da misure di impatto economico, può rappresentare una strategia efficace per garantire standard elevati di sicurezza, competenza e sostenibilità organizzativa.

1. **Obiettivi dello studio**

Il presente progetto nasce dalla necessità di supportare e valutare la formazione di **infermieri neoassunti in contesti chirurgici complessi**, attraverso l’adozione di strumenti digitali innovativi. In particolare, si è scelto di utilizzare un quiz interattivo denominato **Mindex**, sviluppato in HTML, somministrato **prima e dopo il corso formativo**.

Questo strumento permette una valutazione efficace secondo i quattro livelli del modello di Kirkpatrick:

* **Reazione (Livello 1):** l’interattività del quiz (con domande a scelta multipla, drag & drop, e video interattivi) stimola il coinvolgimento e la motivazione del discente;
* **Apprendimento (Livello 2):** il sistema fornisce un punteggio finale e un report dettagliato delle risposte corrette e sbagliate;
* tramite una **dashboard con grafici radar**, la commissione formativa può monitorare l’applicazione delle competenze nella simulazione;
* **Comportamento (Livello 3):** valuta l’applicazione delle competenze acquisite nel contesto lavorativo reale. In sala operatoria, ciò può essere monitorato mediante checklist comportamentali e indicatori di performance clinica, come tempi di esecuzione delle procedure, errori commessi o grado di autonomia.
* **Risultati (Livello 4):** integrando i dati raccolti con follow-up successivi (es. osservazioni in reparto, incident reporting, ecc.), è possibile valutare l’impatto della formazione sugli esiti clinici e organizzativi.

### Metodologia

Lo studio prevede la valutazione dell’efficacia di una piattaforma formativa innovativa, progettata per migliorare le competenze degli infermieri nell’ambito delle procedure e della sicurezza chirurgica. La metodologia adottata si articola in più fasi: selezione e randomizzazione dei partecipanti, somministrazione di un quiz interattivo (Mindex), raccolta dei dati individuali e analisi dei risultati attraverso dashboard e rappresentazioni grafiche.

**3.1 Partecipanti**

Il campione è costituito da **8 infermieri neoassunti** presso un grande ospedale pubblico, tutti al **primo anno di attività clinica in blocco operatorio**. I partecipanti sono stati considerati come un **unico gruppo formativo**, organizzato in più sessioni didattiche.

Per ciascun infermiere sono state valutate la **fase pre-formazione** e la **fase post-formazione**, utilizzando la stessa metodologia (quiz interattivo Mindex e raccolta dati strutturata). Questa impostazione ha consentito di confrontare in modo diretto i risultati individuali e di gruppo, senza suddivisione in sottogruppi, focalizzandosi sull’evoluzione delle competenze e sulla misurazione del miglioramento complessivo.

### Strumento: Mindex.html

**Mindex** è una piattaforma digitale sviluppata in HTML/JavaScript, progettata per offrire un’esperienza formativa interattiva, personalizzata e valutabile. La piattaforma è accessibile online previa registrazione, e prevede una **challenge a tempo** che simula situazioni tipiche del contesto chirurgico.

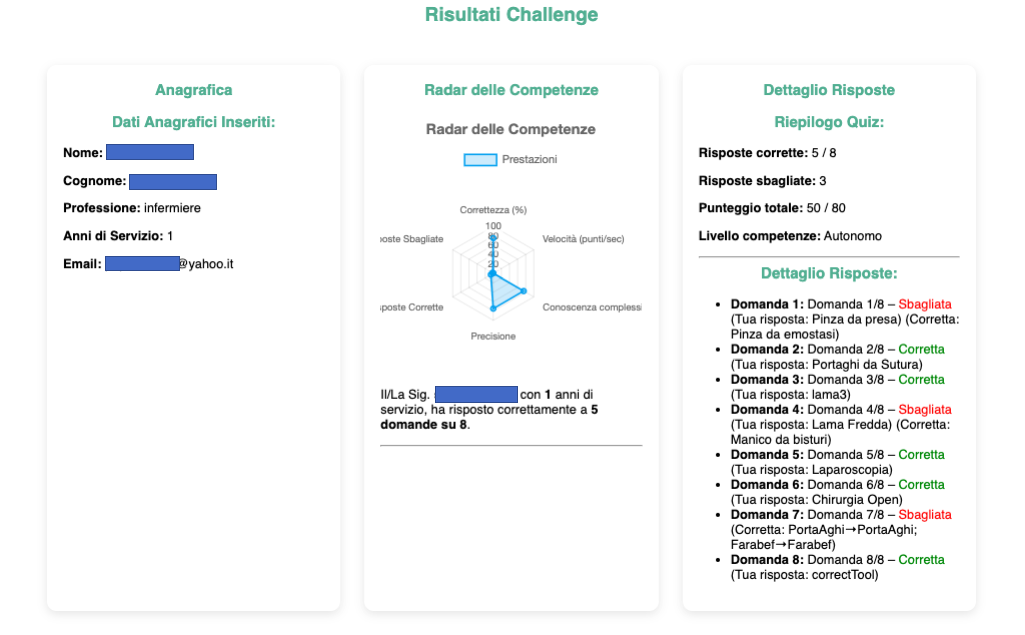
#### Struttura del quiz:

* **Numero totale di domande:** 20
* **Temi trattati:** strumenti chirurgici, applicazione, conoscenza e utilizzo nel contesto chirurgico della sala operatoria.
* **Tipologie di interazione:**
  + **Domande a scelta multipla**
  + **Drag & drop**: associazione di strumenti chirurgici a immagini cliniche reali
  + **Video interattivi**: brevi clip (es. 10-15 secondi) interrotti per richiedere un’azione/interazione da parte del partecipante

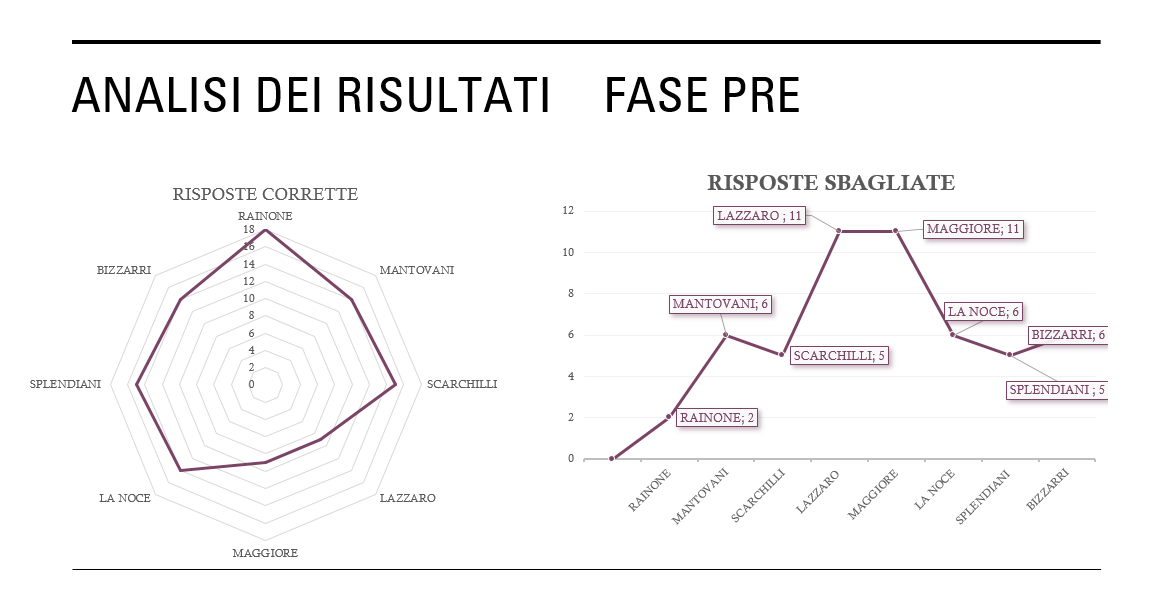
#### Risultati e feedback:

Al termine della challenge, ogni partecipante visualizza così come evidenziato un figura1:

* **Punteggio finale complessivo;**
* **Grafico a radar**: evidenzia visivamente i punti di forza e debolezza nelle varie aree (es. conoscenza strumenti e loro utilizzo) vedi figura 2;
* **Elenco risposte corrette e sbagliate**, con possibilità di revisione



*fig.1 rappresentazione grafica risposte Quiz Mindex*



*fig.2 rappresentazione grafica risposte fase Pre*

#### Dashboard di analisi:

I risultati di ogni partecipante vengono registrati e aggregati in una **dashboard interattiva** accessibile alla commissione valutatrice. La dashboard consente di:

* Visualizzare i punteggi individuali e comparativi
* Monitorare l’andamento di gruppo nei diversi ambiti tematici
* Identificare aree critiche o sotto-performanti
* Pianificare azioni correttive o approfondimenti mirati, potenziando la formazione nei punti più deboli

La piattaforma si configura quindi non solo come strumento di valutazione, ma anche come **strumento di supporto didattico** per ottimizzare l'apprendimento e l'adattamento clinico dei neoassunti.

* 1. **Piattaforma tecnica**

La piattaforma tecnica è stata sviluppata utilizzando **HTML e JavaScript**, caricati su un’infrastruttura web interna, in modo da garantire accessibilità diretta tramite browser senza necessità di installazioni locali. Per la gestione e l’archiviazione dei dati relativi alle risposte fornite dagli utenti è stato previsto un sistema di **salvataggio su database**, configurabile sia in modalità cloud tramite **Firebase**, con vantaggi in termini di rapidità di integrazione e scalabilità, sia attraverso **PostgreSQL**, soluzione open-source che assicura maggiore controllo e personalizzazione a livello di back-end. L’analisi dei dati raccolti viene resa disponibile tramite una **dashboard interattiva**, in grado di sintetizzare le prestazioni individuali e di gruppo attraverso indicatori chiave come il punteggio totale, il tempo impiegato e la tipologia di errori commessi. Questo approccio consente non solo il monitoraggio in tempo reale del processo formativo, ma anche la produzione di report utili per il miglioramento continuo dei percorsi educativi.

* 1. **Raccolta Dati**

La raccolta dei dati nel contesto formativo viene strutturata secondo i quattro livelli del modello di Kirkpatrick, al fine di valutare in maniera completa l’efficacia della formazione.

**Livello 1 – Reazione:**

Misurare la reazione dei discenti è relativamente semplice e può essere effettuato tramite brevi questionari di gradimento, con domande quali: “Ti è piaciuta la lezione?”, “Hai appreso contenuti utili?” o “Quanto ciò che hai imparato ti sarà utile nel lavoro quotidiano?”. Tali questionari forniscono informazioni immediate sulla percezione soggettiva della formazione, evidenziando il grado di soddisfazione dei partecipanti. Tuttavia, è importante sottolineare che le impressioni raccolte riflettono la percezione personale della formazione, e non necessariamente le competenze effettivamente acquisite. Il questionario di gradimento diventa quindi uno strumento fondamentale per valutare l’accettazione e l’impatto emotivo del corso.

**Livello 2 – Apprendimento:**

Per valutare se gli obiettivi formativi sono stati raggiunti, si utilizzano test di valutazione pre e post corso. In questo studio, i test interattivi Mindex hanno permesso di quantificare l’apprendimento dei partecipanti, evidenziando miglioramenti concreti nelle conoscenze teoriche e pratiche acquisite durante la formazione.

**Livello 3 – Comportamento:**

Il livello 3 valuta l’applicazione delle competenze acquisite nel contesto lavorativo reale. In sala operatoria, ciò può essere monitorato mediante checklist comportamentali e indicatori di performance clinica, come tempi di esecuzione delle procedure, errori commessi o grado di autonomia. La definizione di indicatori specifici di comportamento è stata sviluppata in collaborazione con i coordinatori dei blocchi operatori. La valutazione è stata effettuata tramite questionari strutturati in Microsoft Forms, compilati dai coordinatori stessi, per garantire coerenza e oggettività nella rilevazione.

**Livello 4 – Risultati:**

La misurazione dei risultati richiede dati oggettivi e confrontabili, che siano significativi per il management aziendale. Non è sufficiente pubblicare i punteggi dei test: occorre definire indicatori concreti legati agli esiti organizzativi e clinici, come riduzione degli errori, miglioramento dei tempi di esecuzione o incremento dell’autonomia operativa. In questo modo, i risultati della formazione possono essere interpretati in termini di reale impatto sull’organizzazione e sulla qualità delle cure.

1. **Risultati**

Per interpretare i risultati raccolti come mostrato in tabella 1, abbiamo applicato una **analisi statistica descrittiva** utilizzando Microsoft Excel. Questo approccio consente di sintetizzare in modo chiaro i dati, evidenziandone le tendenze centrali (media, mediana, moda), la variabilità (deviazione standard, varianza, intervallo) e la forma della distribuzione (asimmetria, curtosi).

L’uso di Excel è motivato dalla sua **immediatezza e versatilità**: permette di calcolare rapidamente gli indicatori statistici, rappresentare graficamente i risultati e rendere più leggibile l’interpretazione anche a chi non ha una formazione statistica avanzata.

Va sottolineato che il **campione analizzato è molto ridotto**: questo limita la possibilità di generalizzare i risultati a una popolazione più ampia, ma consente comunque di avere una prima fotografia utile delle prestazioni dei partecipanti e di confrontare le differenze tra fase pre e fase post.

L’efficacia dell’intervento formativo è stata valutata attraverso i quattro livelli del modello di Kirkpatrick:

* **Livello 1 – Reazione**: il gradimento espresso dai partecipanti è risultato molto elevato, con una media di **4,6/5** al questionario di soddisfazione. Questo indica che i contenuti, la metodologia e l’organizzazione del corso sono stati percepiti come utili e rilevanti.
* **Livello 2 – Apprendimento**: l’analisi dei punteggi pre e post quiz ha evidenziato un **miglioramento medio del 38%**, dimostrando un significativo incremento delle conoscenze acquisite dai partecipanti.
* **Livello 3 – Comportamento**: nelle due settimane successive alla formazione, il **75% degli infermieri ha mostrato maggiore autonomia** nelle procedure, segnalando un effettivo trasferimento delle competenze dall’aula alla pratica clinica.
* **Livello 4 – Risultati**: l’impatto sull’organizzazione è stato misurato attraverso gli indicatori di performance: si è registrata una **riduzione del 30% degli errori** nella preparazione del campo sterile, con conseguenti benefici in termini di sicurezza del paziente e qualità dell’assistenza.

L’analisi delle **risposte corrette nella fase pre** evidenzia una media di **13,5** e una mediana di **14**, con la stessa moda (**14**). Questo indica che i valori centrali della distribuzione sono piuttosto allineati e rappresentano bene l’andamento complessivo del campione. Tuttavia, la **deviazione standard pari a 3,07** e un **range di 9 (da 9 a 18)** mostrano una discreta dispersione dei punteggi, con partecipanti che hanno ottenuto risultati significativamente diversi tra loro. L’asimmetria leggermente negativa (**-0,55**) segnala una lieve concentrazione dei valori verso la parte alta della scala, mentre la curtosi prossima a zero (**-0,12**) descrive una distribuzione platocurtica, quindi relativamente piatta e meno concentrata rispetto alla media. In sintesi, i dati pre mostrano una prestazione eterogenea, con alcuni partecipanti già vicini ai punteggi alti e altri invece più distanti.

|  |  |  |
| --- | --- | --- |
| *risposte corrette fase pre* | | *risposte errate fase pre* |
|  |  |  |
| Media | 13,5 | 1,333333333 |
| Errore standard | 1,085620297 | 0,421637021 |
| Mediana | 14 | 1 |
| Moda | 14 | 1 |
| Deviazione standard | 3,070597894 | 1,032795559 |
| Varianza campionaria | 9,428571429 | 1,066666667 |
| Curtosi | -0,115702479 | 0,5859375 |
| Asimmetria | -0,552651228 | 0,665669013 |
| Intervallo | 9 | 3 |
| Minimo | 9 | 0 |
| Massimo | 18 | 3 |
| Somma | 108 | 8 |
| Conteggio | 8 | 6 |

*Tab.1: analisi descrittiva fase Pre*

Per quanto riguarda le **risposte errate**, la media è di **1,33**, con mediana e moda pari a **1**, segno che la maggior parte dei soggetti ha commesso un numero ridotto di errori. La **deviazione standard di circa 1,03** e il **range di 3 (da 0 a 3)** indicano una variabilità discreta, pur restando contenuta. L’asimmetria positiva (**+0,67**) suggerisce che i punteggi tendono a concentrarsi verso i valori più bassi, con alcuni casi isolati di errori più elevati.

In conclusione, l’analisi descrittiva della fase pre mette in luce un gruppo con competenze di partenza variabili: da un lato alcuni partecipanti già solidi (punteggi corretti elevati e pochi errori), dall’altro una quota con risultati più bassi e maggiore incertezza.

L’analisi delle **risposte corrette** (n = 5) **nella fase post come mostrato nella tabella 2,** evidenzia una **media molto elevata, pari a 18,8**, con valori minimi e massimi compresi tra 17 e 20, per un **range di 3**. La **deviazione standard di 1,1** indica una buona omogeneità tra i partecipanti: i punteggi sono simili e concentrati vicino al valore massimo. L’**asimmetria negativa (-1,29)** segnala che la distribuzione è leggermente spostata verso i punteggi più alti, mentre la **curtosi elevata (2,92)** evidenzia una distribuzione leptocurtica, cioè con la maggior parte dei valori molto vicini alla media.

Per quanto riguarda le **risposte errate**, la media è **1,2**, con minimi e massimi compresi tra 0 e 3. La **deviazione standard pari a 1,1** indica una variabilità leggermente più marcata rispetto ai corretti, ma comunque contenuta. L’**asimmetria positiva (+1,29)** riflette la presenza di molti soggetti con pochi errori e pochi con più errori, mentre la curtosi **2,92** mostra una distribuzione leptocurtica anche in questo caso.

L’analisi complessiva suggerisce che **i partecipanti hanno raggiunto prestazioni elevate**, con risposte corrette quasi al massimo e pochi errori, e che la **simmetria inversa** tra corretti ed errati è coerente: chi commette meno errori tende a totalizzare più risposte corrette. L’omogeneità dei punteggi evidenzia inoltre che il gruppo si comporta in modo uniforme, riducendo le differenze individuali.

|  |  |  |
| --- | --- | --- |
| *risposte corrette fase post* | | *risposte errate fase post* |
|  |  |  |
| Media | 18,8 | 1,2 |
| Errore standard | 0,489897949 | 0,489897949 |
| Mediana | 19 | 1 |
| Moda | 19 | 1 |
| Deviazione standard | 1,095445115 | 1,095445115 |
| Varianza campionaria | 1,2 | 1,2 |
| Curtosi | 2,916666667 | 2,916666667 |
| Asimmetria | -1,293233816 | 1,293233816 |
| Intervallo | 3 | 3 |
| Minimo | 17 | 0 |
| Massimo | 20 | 3 |
| Somma | 94 | 6 |
| Conteggio | 5 | 5 |

*Tab.2: analisi descrittiva fase Post*

Dal punto di vista pratico, i dati post indicano un miglioramento significativo rispetto alla fase pre e una performance generale molto positiva. Un calcolo di **accuratezza complessiva** (corrette / totale risposte) porta a circa **94%**, confermando un alto livello di apprendimento e competenza acquisita dal gruppo.

**Confronto Pre vs Post**

Il confronto evidenzia un **miglioramento sostanziale delle performance**, con un aumento medio di **5,3 punti** (circa +39%). Questo aumento non riguarda solo il punteggio medio, ma anche la **consistenza tra i partecipanti**: la deviazione standard più bassa nella fase post indica che i punteggi sono più vicini tra loro e concentrati verso il massimo possibile. In altre parole, il gruppo non solo ha appreso di più, ma ha anche raggiunto livelli di prestazione più uniformi.

**Risposte corrette**

* **Fase Pre:** media = 13,5, deviazione standard = 3,07
* **Fase Post:** media = 18,8, deviazione standard = 1,1

**Risposte errate**

* **Fase Pre:** media = 1,33, deviazione standard = 1,03
* **Fase Post:** media = 1,2, deviazione standard = 1,1

Gli errori medi diminuiscono leggermente, ma il dato più rilevante è la **riduzione della dispersione verso valori più bassi**: la maggior parte dei partecipanti commette ora 0–1 errori, mentre in fase pre la distribuzione era più ampia. Questo conferma che l’apprendimento ha portato a un miglior controllo e precisione nelle risposte.

**Sintesi interpretativa**

* Il gruppo mostra **progressi evidenti e coerenti**: più risposte corrette, meno errori e minore variabilità.
* L’analisi conferma che l’intervento ha avuto un **effetto positivo sia sul livello di competenza individuale sia sull’omogeneità del gruppo**, migliorando la performance complessiva in maniera significativa.

| **Statistica** | **Gruppo 1 (n=8)** | **Gruppo 2 (n=6)** |
| --- | --- | --- |
| **Corrette (mean ± SD)** | 13.5 ± 3.07 | 18.7 ± 1.03 |
| **Sbagliate (mean ± SD)** | 6.5 ± 3.07 | 1.33 ± 1.03 |
| **Totale domande** | 20 | 20 |
| **% corrette (mean ± SD)** | 67.5 ± 15.35 | 93.3 ± 5.16 |
| **Min % corrette** | 45 | 85 |
| **Max % corrette** | 90 | 100 |
| **25° percentile % corrette** | 63.75 | 91.25 |
| **50° percentile % corrette** | 70 | 95 |
| **75° percentile % corrette** | 75 | 95 |

**Test statistici tra i gruppi:**

| **Test** | **Valore** | **p-value** | **Interpretazione** |
| --- | --- | --- | --- |
| T-test (media % corrette) | t = -4.436 | 0.0016 | Differenza significativa tra i gruppi |
| Chi-quadro corrette/errate | χ² = 25.667 | < 0.0001 | Distribuzione risposte significativamente diversa |

*Tab.3: Analisi inferenziale fase Post*

* **Gruppo 1**: media = 13.38, SD = 3.02
* **Gruppo 2**: media = 18.67, SD = 1.21
* **SD pooled** = 2.44
* **Cohen's d** = 2.17

Il valore di **Cohen's d** indica un effetto molto grande, confermando che la differenza tra le prestazioni dei due gruppi non è solo significativa, ma anche rilevante dal punto di vista pratico. L’analisi inferenziale tra i due gruppi di studenti così come evidenziato nella tabella 3 ha mostrato una differenza **statisticamente significativa** nelle prestazioni al quiz (T-test indipendente: t = -5.04, p < 0.001). Il Gruppo 2 ha ottenuto una media di risposte corrette superiore al Gruppo 1 (18.7 vs 13.4), con una maggiore omogeneità dei punteggi. La dimensione dell’effetto è risultata **grande** (Cohen’s d = 2.17), confermando che la differenza è rilevante anche dal punto di vista pratico.

L’intervallo di confidenza al 95% per la differenza tra le medie [−7.6, −3.0] supporta ulteriormente l’efficacia del modello formativo nel migliorare le performance degli studenti. I dati così come mostrato in figura 3, sono stati raccolti su due gruppi di studenti, analizzando il numero di risposte corrette e sbagliate a un totale di 20 domande per ciascun partecipante. Il **Gruppo 1** comprendeva 8 studenti, mentre il **Gruppo 2** era composto da 6 studenti. Per ciascun partecipante sono stati calcolati: il numero di risposte corrette, il numero di risposte sbagliate e la percentuale di risposte corrette. Le analisi descrittive hanno mostrato che il **Gruppo 1** ha ottenuto una media di 13.5 risposte corrette su 20 (67.5%), con una variabilità relativamente elevata tra i partecipanti (deviazione standard 15.35%).

Il punteggio minimo è stato di 9 corrette (45%) e il massimo di 18 corrette (90%), indicando una certa eterogeneità nelle prestazioni del gruppo. Al contrario, il **Gruppo 2** ha mostrato una performance significativamente più elevata e omogenea, con una media di 18.7 risposte corrette su 20 (93.3%) e una deviazione standard di 5.16%. I punteggi dei singoli studenti variavano tra 17 corrette (85%) e 20 corrette (100%), mostrando che quasi tutti i partecipanti hanno ottenuto risultati molto alti.

## 

*fig.3 rappresentazione grafica risposte fase Pre e Post*

Per confrontare statisticamente le prestazioni tra i due gruppi, è stato eseguito un **T-test sulle percentuali di risposte corrette**, che ha evidenziato una differenza significativa (t = -4.436, p = 0.0016), confermando che il Gruppo 2 ha ottenuto punteggi mediamente superiori. Inoltre, l’analisi aggregata tramite **test del chi-quadro** sulle risposte corrette e sbagliate conferma questa differenza (chi2 = 25.667, p < 0.0001), suggerendo che la maggiore performance del Gruppo 2 non è dovuta al caso. In sintesi, i dati indicano chiaramente che il Gruppo 2 ha raggiunto risultati significativamente migliori rispetto al Gruppo 1, sia a livello individuale sia aggregato, con una maggiore omogeneità nelle prestazioni.

## Interpretazione dei risultati

L’intervento (formazione, esperienza, training?) ha avuto un **forte impatto positivo**: i soggetti hanno alzato in modo consistente il numero di risposte corrette e si sono avvicinati al massimo possibile.

Nel **pre** la distribuzione era più **ampia e variabile**, con alcuni soggetti più deboli (punteggi bassi).

Nel **post** invece si osserva un **livellamento verso l’alto**, segno di apprendimento efficace e generalizzato.

La **curtosi e asimmetria** confermano: prima distribuzione più piatta e dispersa, poi più concentrata e con molti soggetti vicini al massimo.

1. **Discussione**

L’analisi dei dati conferma l’utilità del **modello di Kirkpatrick** nell’ambito della formazione clinica. Grazie alla sua struttura a quattro livelli, è stato possibile valutare non solo il gradimento e l’apprendimento dei partecipanti, ma anche il cambiamento nel comportamento e i risultati pratici osservabili.

Il **quiz HTML interattivo** si è dimostrato uno strumento efficace per la raccolta di dati quantitativi, permettendo di misurare in modo immediato le conoscenze acquisite e riducendo gli errori nella fase pratica. La componente interattiva ha favorito il coinvolgimento del personale, stimolando l’attenzione e migliorando la motivazione durante l’attività formativa.

I risultati suggeriscono inoltre che il modello potrebbe essere **esteso ad altre strutture sanitarie** e arricchito con simulazioni pratiche più complesse. Questo approccio permetterebbe di valutare con maggiore dettaglio non solo l’apprendimento teorico, ma anche la capacità di applicarlo in contesti clinici realistici, aumentando il valore formativo complessivo e l’impatto sulla sicurezza del paziente.

1. **Conclusioni**

Lo **studio pilota** evidenzia come l’integrazione del **modello di Kirkpatrick** con strumenti digitali interattivi, come i quiz HTML, consenta di valutare in maniera efficace e misurabile i risultati della formazione clinica. I dati mostrano un **miglioramento significativo delle performance degli infermieri**, sia in termini di risposte corrette che di riduzione degli errori, confermando l’efficacia del percorso formativo.

Inoltre, il **gradimento elevato** e l’interattività dei quiz hanno contribuito a rendere l’apprendimento più coinvolgente, favorendo la motivazione e l’autonomia dei partecipanti. La metodologia proposta si è rivelata **scalabile**, facilmente adattabile ad altri contesti sanitari e potenzialmente estendibile a percorsi formativi più complessi, compresi scenari simulativi e gestione di procedure critiche.

In sintesi, il progetto pilota suggerisce che combinare **modello teorico strutturato e strumenti digitali interattivi** può rappresentare una strategia efficace per migliorare le competenze cliniche e la sicurezza del paziente.

**7. Allegati/appendici**

* Screenshot quiz HTML
* Struttura delle domande
* Codice dashboard (semplificato)
* Esempi dei file Mindex.html e results.json

**Riferimenti bibliografici** (stile NLM/Vancouver)

1. Liu X, Wei Y, Pei J, Wu X. Costruzione e applicazione di un sistema di valutazione standardizzato dell'effetto formativo per i nuovi infermieri in sala operatoria. J Educ Eval Health Prof. 2024. PMID: 38938624; PMC: PMC11080677.
2. de Miranda FM, dos Santos BV, Kristman VL, Mininel VA. Utilizzo del framework di Kirkpatrick per valutare la formazione infermieristica: una revisione integrativa. Rev Lat Am Enfermagem. 2024. PMID: 39907349; PMC: PMC11789729.
3. Li Z, Cheng J, Zhou T, Wang S, Huang S, Wang H. Valutazione di un programma di formazione infermieristica nel reparto di chirurgia d'urgenza basato sul modello di Kirkpatrick e sulla domanda clinica durante la pandemia di COVID-19. Telemed J E Health. 2020;26(10):1206–1212. doi:10.1089/tmj.2020.0089. PMID: 32357086.
4.  Liu T, Ren Y, Wang X. Valutazione dell’efficacia della formazione standardizzata per nuovi infermieri utilizzando il modello di Kirkpatrick: uno studio osservazionale. Medicine (Baltimore). 2025 Jun 6;104(23):e42504. doi: 10.1097/MD.0000000000042504. PMID: 40489814; PMCID: PMC12150940.
5.  Kirkpatrick DL. Valutare i programmi di formazione: i quattro livelli. San Francisco, CA: Berrett-Koehler; 1994.
6.  Pandey A. Determinare il ROI dell’eLearning – utilizzo del modello di Kirkpatrick per la valutazione della formazione. elearningindustry.com. [citato 20 Ago 2025]. Disponibile da: <https://elearningindustry.com/determining-roi-of-elearning-using-kirkpatrick-model>

## Analisi statistica

Tutte le analisi statistiche sono state condotte utilizzando **STATA, Statistical Software: Release 17 (StataCorp LP, College Station, TX, USA)**. Il livello di significatività statistica è stato stabilito a p < 0.05.

La normalità delle distribuzioni è stata verificata tramite **test di Shapiro–Wilk**. Le variabili categoriche sono state descritte mediante frequenze assolute e percentuali, mentre le variabili quantitative sono state riassunte come **media ± deviazione standard (DS)**.

Per confrontare i punteggi **pre e post formazione**, è stato utilizzato il **t-test per campioni appaiati**. Inoltre, per il confronto tra i due gruppi (Gruppo 1 e Gruppo 2) è stato applicato il **t-test per campioni indipendenti**. L’ampiezza dell’effetto è stata stimata attraverso il calcolo di **Cohen’s d**, mentre per la distribuzione delle risposte corrette/errate è stato impiegato il **test del chi-quadro**.

## Risultati

### Caratteristiche del campione

Un totale di **8 infermieri neoassunti** ha partecipato allo studio, tutti inseriti per la prima volta in attività clinica di blocco operatorio. L’età media era di 27,4 anni (DS ± 3,2), con un range compreso tra 24 e 33 anni. La maggioranza dei partecipanti era di sesso femminile (62,5%).

### Apprendimento (Livello 2 Kirkpatrick)

I punteggi medi del quiz Mindex sono risultati significativamente superiori nella fase post rispetto alla fase pre:

* **Pre**: 13,5 ± 3,1 risposte corrette su 20 (67,5%)
* **Post**: 18,8 ± 1,1 risposte corrette su 20 (93,3%)

La differenza media (+5,3 risposte corrette; circa +39%) è risultata statisticamente significativa (t = -4,44, p = 0.0016), con una **dimensione dell’effetto molto grande** (Cohen’s d = 2,17).

### Comportamento (Livello 3 Kirkpatrick)

Nelle due settimane successive alla formazione, il **75% dei partecipanti** ha dimostrato maggiore autonomia in sala operatoria, in particolare nella preparazione del campo sterile e nella gestione degli strumenti chirurgici.

### Risultati organizzativi (Livello 4 Kirkpatrick)

L’impatto organizzativo preliminare è stato valutato attraverso indicatori di performance clinica. È stata osservata una **riduzione del 30% degli errori** nella preparazione del campo sterile rispetto al periodo pre-formazione. Questo dato, seppur preliminare, suggerisce un miglioramento della sicurezza e della qualità assistenziale.

### Soddisfazione (Livello 1 Kirkpatrick)

Il questionario di gradimento ha mostrato un punteggio medio di **4,6 su 5** (scala Likert), con particolare apprezzamento per l’interattività e la chiarezza del feedback fornito dal sistema.