Bollettino Unione Matematica Italiana

Sezione A – La Matematica nella Società e nella Cultura

Alan H. Schoenfeld

Obiettivi e metodi di ricerca in didattica della matematica

Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. **3-A**—*La Matematica nella Società e nella Cultura* (2000), n.2, p. 175–199.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2000_8_3A_2_175_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2000.

Bollettino U. M. I.
La Matematica nella Società e nella Cultura
Serie VIII, Vol III-A, Agosto 2000, 175-199

Obiettivi e metodi di ricerca in didattica della matematica (*).

ALAN H. SCHOENFELD

Bertrand Russel ha definito la matematica come la scienza in cui non sappiamo mai di cosa stiamo parlando o se ciò che stiamo dicendo è vero. Si è visto che la matematica è ampiamente utilizzata in altri campi della scienza. Dunque, la maggior parte degli altri scienziati non sa di cosa sta parlando o se quello che sta dicendo è vero.
Joel Cohen, «On the nature of mathematical proofs»

Non ci sono dimostrazioni in didattica della matematica.
Henry Pollak

La prima citazione qui sopra è spiritosa; la seconda è seria. Entrambe servono però a mettere in luce alcune delle maggiori differenze tra la matematica e la didattica della matematica — differenze che devono essere tenute in considerazione se si vuole capire la natura di metodi e risultati nel campo della didattica della matematica.

La citazione di Cohen indica alcuni aspetti seri della matematica. Per esempio, parlando delle varie geometrie si comincia con i termini primitivi (indefiniti). Quindi, applicando le regole della logica, si dimostra che se qualcosa è vero allora altri risultati devono necessariamente seguire. Da un lato, i termini primitivi sono indefiniti, ovvero «non sappiamo di cosa stiamo parlando»; dall'altro lato, i risultati sono definitivi. Come avrebbe potuto dire Gertrude Stein, una dimostrazione è una dimostrazione è una dimostrazione.

(*) Traduzione dell'articolo di Alan H. Schoenfeld, *Purposes and Methods of Research in Mathematics Education*, apparso sulle Notices Amer. Math. Soc. **47**, 6, 641-649. Ringraziamo il prof. A. Schoenfeld e l'American Mathematical Society di averci permesso traduzione e pubblicazione di questo articolo. La traduzione è dovuta a Federica Olivero.

Altre discipline funzionano in modo diverso. La frase di Pollak (che è un commento alla conferenza «Cognitive Science and Mathematics Education» 1984) non vuole essere una svalutazione della didattica della matematica, bensì vuole puntare l'attenzione sul fatto che la natura delle prove e dei ragionamenti nel campo della didattica della matematica è abbastanza diversa dalla natura delle prove e dei ragionamenti nel campo della matematica. Sicuramente le domande che ci si può porre (e a cui ci si aspetta di poter dare una risposta) nel campo della ricerca didattica non sono le stesse domande che si possono porre i matematici. Inoltre, i matematici e i ricercatori in didattica tendono ad avere prospettive diverse per quanto riguarda gli obiettivi e gli scopi della ricerca in didattica della matematica.

Questo articolo comincia con un tentativo di presentare alcune delle prospettive più importanti e fornire una base generale riguardante la natura della ricerca in didattica della matematica. Tra le questioni affrontate ci sono le seguenti: quali sono gli obiettivi della ricerca in didattica della matematica? Quali sono i modelli e le teorie prodotte nel campo della didattica, confrontati con quelli prodotti in matematica e nel campo delle scienze fisiche? A quali tipi di domande la ricerca didattica può rispondere? Date queste domande, quali ne sono le risposte plausibili? Quale evidenza è necessaria per sostenere le affermazioni riguardanti la didattica? Quali metodi possono generare questa evidenza? Quali potrebbero essere gli standard per valutare affermazioni, modelli e teorie? Come vedremo in seguito, le risposte a queste domande differiscono sostanzialmente a seconda che vengano affrontate dal punto di vista della matematica o della didattica.

Obiettivi.

La ricerca in didattica della matematica ha due obiettivi principali, uno teorico e uno applicativo:

- puro (scienza di base): capire la natura del modo di pensare, insegnare ed apprendere in matematica;

- applicativo (ingegneria didattica): utilizzare queste conoscenze per migliorare l'insegnamento della matematica.

Questi obiettivi sono strettamente collegati e il primo è almeno tanto importante quanto il secondo. Il motivo è semplice: senza comprendere in profondità come pensare, insegnare ed apprendere non è possibile alcun progresso per quanto riguarda il «fronte applicativo». Un'utile analogia è quella con la ricerca e la pratica nel campo medico. Esiste una grande varietà di tipi di ricerca in campo medico. Alcune ricerche sono urgenti e hanno possibili applicazioni nel futuro prossimo. Altri tipi di ricerca hanno come obiettivo il comprendere i meccanismi fisiologici di base. A lungo termine questi due tipi di lavori sono in sinergia. Questo *sia* perché le conoscenze di base sono intrinsecamente interessanti *sia* perché esse costituiscono e rinforzano le fondamenta su cui si basa il lavoro pratico.

È necessario comprendere a fondo questi obiettivi duali. Essi sono molto diversi dall'unico obiettivo che viene solitamente attribuito alla ricerca in didattica della matematica da molti matematici:

- «Dimmi cosa funziona in classe.»

Questo non vuol dire che i matematici non siano interessati alla ricerca di base in didattica della matematica, però il loro interesse principale riguarda l'utilità in termini diretti. Sicuramente la comunità didattica deve fornire risultati utili — di certo l'utilità costituisce una motivazione per la grande maggioranza dei lavori nel campo della ricerca didattica — però è un errore pensare che le applicazioni dirette (curriculum, sviluppo «dimostrazione» che un trattamento didattico funziona, etc.) siano l'obiettivo principale della ricerca in didattica della matematica.

Domande.

Una questione importante che deve essere discussa quando si pensa a che cosa la didattica della matematica possa offrire è la seguente: a quali tipi di domande la ricerca in didattica della matematica può rispondere? Le domande più comunemente formulate dai matematici — «Che cosa funziona?», «Qual è l'approccio migliore?» — in linea di principio non possono trovare risposta. Il motivo è che ciò che una persona ritiene che funzioni dipende da che cosa quella persona considera importante. Prima di provare a stabilire se un approccio didattico ha successo, è necessario porsi delle domande, come: che cosa si vuole ottenere? Quale comprensione, per quali studenti, in che condizioni, con quali limitazioni? Consideriamo gli esempi seguenti.

Una domanda che viene frequentemente posta dai dirigenti scolastici è: «I corsi numerosi sono validi come quelli meno numerosi?». Spero che sia chiaro che una tale domanda non può trovare risposta in termini astratti. Quanto essere soddisfatti di corsi numerosi dipende da quali obiettivi si ritengono importanti. Quanto conta il coinvolgimento degli studenti? È importante cosa pensano gli studenti del corso e del dipartimento? È importante il numero di studenti che in seguito decidono di continuare a studiare matematica? La conclusione che si può trarre riguardo all'utilità delle classi numerose può variare molto, a seconda di quanto peso viene dato a questi obiettivi.

Simili argomenti compaiono anche se si considera solo la matematica che viene insegnata. Supponiamo che ci si ponga la seguente domanda: gli studenti di corsi numerosi imparano tanta matematica quanto gli studenti di corsi meno numerosi? Immediatamente ci si deve chiedere: «che cosa conta come matematica? quanto peso viene dato (per esempio) al problem solving, alla modellizzazione, o alla capacità di comunicare in matematica?» I giudizi riguardanti l'efficacia di una forma di insegnamento piuttosto che un'altra dipendono da quali risposte si danno a queste domande. Per farla breve, un ricercatore deve sapere che cosa cercare e che cosa considerare come prova di quello che cerca, prima di poter dire se effettivamente c'è.

Il fatto che i nostri giudizi riflettano i nostri valori va considerato anche quando si affrontano domande del tipo: quale approccio funziona meglio (o addirittura è il migliore)? Questo può sembrare ovvio, ma spesso non lo è. Consideriamo la riforma del «Calculus». Subito dopo la conferenza «Lean and Lively», i cui atti sono stati pubblicati in Douglas [5], la National Science Foundation (NSF) ha finanziato un'importante iniziativa per la riforma del «Calculus». A metà degli anni '90 i componenti della commissione della NSF erano convinti che la riforma del «Calculus» fosse una «cosa buona» e che essa avrebbe dovuto essere un modello per tutte le riforme anche riguardanti altri contenuti. La NSF convocò alcuni matematici che erano stati coinvolti nella riforma e alcuni ricercatori in didattica della matematica e pose loro la seguente domanda: «Abbiamo qualche prova che mostri che la riforma del «Calculus» ha funzionato? (ovvero che il «Calculus» insegnato col metodo della riforma sia migliore del «Calculus» tradizionale?)». Si pensava sostanzialmente a qualche forma di test. Pensavano che sarebbe stato facile preparare un test, presentarlo agli studenti e mostrare quindi che gli studenti della riforma avevano avuto risultati migliori.

Coloro che suggerivano questo approccio non avevano capito che quello che proponevano era un confronto tra mele e arance. Se si proponesse un test tradizionale basato sull'abilità di fare manipolazioni simboliche, gli studenti della «riforma» sarebbero in svantaggio perché non hanno esercitato le abilità di calcolo. Se si desse un test basato sulla tecnologia o con una forte componente di modellizzazione, gli studenti tradizionali sarebbero svantaggiati perché tecnologie e modellizzazione non sono trattati in modo ampio nel loro programma. In tutti e due i casi dare un test e confrontare i risultati sarebbe scorretto. Un modo appropriato per procedere sarebbe quello di esaminare i programmi, identificare gli argomenti importanti e specificare che cosa vuol dire aver conseguito una comprensione concettuale degli argomenti. Con queste informazioni, i singoli dipartimenti e istituzioni potrebbero poi decidere quali siano gli aspetti più importanti della comprensione che si vogliono valutare e come. Dopo una lunga discussione, l'attenzione della NSF si spostò dalla documentazione degli effetti della riforma del «Calculus» allo sviluppo di un framework per investigare gli effetti dell'insegnamento del «Calculus». Il risultato di questo lavoro fu il libro *Student Assessment in Calculus* [10] pubblicato nel 1997.

Riassumendo, molte delle domande che sarebbe naturale porre — domande del tipo: che cosa funziona? o qual è il metodo migliore? — non possono trovare risposta, per ottime ragioni.

Dunque, quali tipi di questioni la ricerca in didattica della matematica *può* affrontare? Direi che alcuni dei contributi fondamentali della ricerca in didattica della matematica sono i seguenti:

- prospettive teoriche per comprendere il pensare, insegnare ed apprendere;
- descrizione di aspetti cognitivi (per esempio, pensare matematicamente, comprensione e fraintendimenti da parte degli studenti dei concetti di funzione, limite, ecc.);
- dimostrazioni di esistenza (casi in cui gli studenti imparano il problem solving, l'induzione, la teoria dei gruppi; possibilità di attuazione di diversi tipi di istruzione);
- descrizione delle conseguenze (positive o negative) di varie forme di insegnamento.

L'articolo [1] di Michele Artigue descrive molti risultati di tali studi. Io ne descriverò altri e parlerò dei metodi per ottenere tali risultati nella sezione «Metodi».

Teorie e modelli (e criteri per decidere quali siano validi).

Quando i matematici usano i termini «teoria» e «modelli», solitamente hanno in mente delle cose ben precise, sia per quanto riguarda la natura di queste entità, sia per quanto riguarda il tipo di prove necessarie per fare delle affermazioni su quelli. I termini «teoria» e «modelli» a volte sono usati in modo diverso nelle scienze naturali e nelle scienze sociali. In questo paragrafo esaminerò brevemente gli esempi della Tabella 1.

TABELLA 1. – *Teorie e modelli in matematica/fisica, biologia e didattica/psicologia* (**).

Soggetto	Matematica, fisica	Biologia	Didattica, psicologia
Teoria delle/a...	Equazioni; gravità	Evoluzione	Mente
Modello del...	Diffusione del calore in una lamina	Relazioni preda-predatore	Problem solving

(**) Riproduzione autorizzata, tratta da Schoenfeld [11], pag. 9.

In matematica, le teorie sono presentate esplicitamente, come ad esempio la teoria delle equazioni o la teoria delle variabili complesse. I risultati sono ottenuti analiticamente: si dimostra che gli oggetti considerati hanno effettivamente le proprietà ipotizzate. Nella fisica classica c'è un simile grado di specificità; per esempio i fisici specificano una legge dell'inverso del quadrato per l'attrazione gravitazionale. I modelli sono considerati approssimazioni, ma ci si aspetta che siano approssimazioni molto precise in forma deterministica. Quindi, per esempio per modellizzare la diffusione del calore in una lamina, si specificano le condizioni iniziali al bordo e le condizioni della diffusione del calore, e poi si risolvono le equazioni appropriate. Non c'è ambiguità nel processo. Le descrizioni sono esplicite e lo standard per la correttezza è la dimostrazione matematica. Teoria e modelli che ne derivano possono essere utilizzati per fare previsioni, che a loro volta sono prese come validazione empirica della correttezza della teoria.

Le cose sono molto più complesse nelle scienze biologiche. Consideriamo, per esempio, la teoria dell'evoluzione. I biologi sono in generale d'accordo sulla correttezza di questa teoria, però l'evidenza che la sorregge è diversa dal tipo di evidenza che sorregge le teorie in matematica o fisica. Non esiste alcun modo per dimostrare che la teoria dell'evoluzione è corretta in senso matematico; gli argomenti che la sostengono sono (per dirla con le parole di un libro di Polya) «modelli di ragionamento plausibile», insieme all'attenta considerazione di ipotesi alternative. In effetti, i biologi hanno detto questo: «c'è una grande quantità di dati coerenti con la teoria; non esiste alcuna chiara prova che falsifichi la teoria proposta e non ci sono altre ipotesi che soddisfino gli stessi criteri.» Anche se previsioni di eventi futuri non sono possibili, data la scala temporale degli eventi legati all'evoluzione, la teoria consente una forma alternativa di previsione. Reperti fossili non ancora esaminati devono conformarsi alla teoria, cosicchè la teoria possa essere usata per descrivere delle proprietà che i fossili dovrebbero o non dovrebbero avere in particolari strati geologici. La documentazione cumulativa è considerata supporto alla teoria.

In breve, teoria e prova che la sorregge possono essere molto diverse nelle scienze naturali e in matematica e fisica. Lo stesso vale per i modelli, o per lo meno per il grado di precisione che ci si aspetta da questi: non ci si aspetta di certo che il comportamento della popolazione animale sia conforme ai modelli preda-predatore nello stesso modo in cui ci si aspetta che la diffusione del calore in una lamina sia conforme ai modelli di diffusione del calore.

Infine, teorie e modelli nelle scienze sono costantemente soggetti a revisione e raffinamento. Per quanto meravigliosa e gloriosa fosse la teoria gravitazionale di Newton, essa è stata sostituita dalla teoria della relatività di Einstein. Oppure consideriamo la teoria nucleare. La teoria delle valenze, basata su modelli di elettroni che ruotano su orbite intorno ai nuclei, ha permesso di fare incredibili previsioni, come per esempio l'esistenza di elementi ancora da scoprire. Ma i fisici non parlano più di orbite di elettroni attorno ai nuclei; quelle che erano particelle solide nella teoria, come gli elettroni, sono state rimpiazzate da nuvole elettroniche probabilistiche. Le teorie evolvono.

La ricerca in didattica della matematica ha molte caratteristiche simili a quelle della ricerca in fisica e nelle scienze naturali, che sono state descritte sopra. Nella «teoria della mente» per esempio si fanno delle congetture sulla natura dell'organizzazione della mente — per esempio, che ci siano certi tipi di strutture mentali che funzionano in un determinato modo. Una di queste è che esistano diversi tipi di memoria, tra cui la memoria di lavoro o memoria a «breve termine». Secondo questa teoria, l'atto del pensare utilizza questa memoria di lavoro: gli «oggetti del pensiero» che sono manipolati mentalmente vengono immagazzinati temporaneamente nella memoria di lavoro. La cosa interessante (e scientifica) è che questa teoria stabilisce anche i limiti della memoria di lavoro: è stato detto (per esempio in [8]) che le persone non possono immagazzinare più di nove «pezzi» di informazione alla volta nella memoria di lavoro.

Per vedere se tale affermazione è vera, si potrebbe provare a moltiplicare 379 per 658 con gli occhi chiusi. La maggior parte della gente troverà questo difficile o addirittura impossibile. (In un recente meeting ho proposto questo problema a 75 matematici. Nessuno di loro è riuscito a dare una risposta nel giro di pochi minuti). Il motivo è che il numero di cose che si devono ricordare — i numeri originali e i prodotti parziali nella moltiplicazione — è maggiore di nove. Ora, siamo in grado di svolgere meglio questo compito mentalmente dopo aver ritenuto a memoria alcuni dei risultati parziali: per esempio, si potrebbe calcolare $8 \times 379 = 3032$ e ripetere mentalmente «3032» finchè questo diventa un pezzo e occupa solo uno spazio («buffer») nella memoria di lavoro. Questo lascia uno spazio di lavoro sufficiente per fare altri calcoli. Utilizzando questo tipo di spezzettamento, si riesce a trascendere i limiti della memoria di lavoro ([1]).

Ora consideriamo lo stato di verità dell'affermazione che la memoria di lavoro non ha più di nove caselle. Non si potrà mai avere una dimostrazione assoluta di questa affermazione. Primo, è molto improbabile che i ricercatori trovino la locazione fisica delle caselle della memoria di lavoro nel cervello, anche se queste dovessero esistere; le caselle sono componenti di modelli e non sono necessariamente oggetti fisici. Secondo, la prova in favore di questa asserzione è convincente ma non può essere definitiva. Sono stati condotti molti esperimenti in cui vengono assegnati problemi che richiedono l'uso di più di nove caselle della memoria di lavoro, e i soggetti non sono riusciti a risolverli (oppure, dopo un pò di sforzo, li risolvevano portando avanti una qualche specie di spezzettamento).

Come per la teoria dell'evoluzione, c'è una grande quantità di prove che sono coerenti con questa affermazione, non c'è prova che la contraddica e nessuna altra ipotesi risponde agli stessi criteri. Pe-

([1]) Il meccanismo di «spezzettare» viene usato in molti contesti. Un esempio banale: ci si può ricordare di un numero di telefono di 10 cifre memorizzando blocchi di 3 cifre come unità. La teoria afferma poi che lo spezzettare è ciò che ci permette di leggere questo articolo. Ogni parola che leggiamo è un pezzo, che prima era un insieme di lettere che dovevano essere lette separatamente. La stessa cosa succede con tutti i concetti matematici che teniamo a mente come un'unità. Infine, i «brillanti calcolatori» — le persone che riescono a fare mentalmente dei grandi calcoli molto rapidamente — sono un controesempio a questa teoria? Non sembra essere così. Le persone che sono state studiate sembravano aver memorizzato un grande numero di risultati intermedi. Per esempio, molti richiameranno alla mente direttamente 72 come un «pezzo» quando si tratta di fare un calcolo che coinvolge 9×8; i «brillanti calcolatori» probabilmente fanno la stessa cosa con numeri di 2 o 3 cifre. Questo riduce il carico della memoria di lavoro.

rò questa affermazione è dimostrata? No, almeno non in senso matematico. Lo standard appropriato è in sostanza ciò che una giuria considererebbe evidenza al di là di ogni ragionevole dubbio. La stessa cosa succede con altri modelli, come quelli che riguardano il «problem solving» o (il mio attuale interesse) i modelli di insegnamento (cf. [12], [13]). Sto attualmente cercando di costruire una descrizione teorica che spieghi come e perché gli insegnanti fanno quello che fanno in classe. Questo lavoro, elaborato con lo stesso livello di dettagli di una teoria sulla memoria, è chiamato «teoria dell'insegnamento contestualizzato». Il punto è che grazie alla teoria e avendo tempo sufficiente per modellizzare un insegnante particolare, si potrebbe costruire una descrizione del modo di insegnare di quella persona che caratterizzi il suo comportamento in classe in modo molto preciso. Quando si considera questo tipo di lavoro, non ci si può aspettare di trovare la precisione propria della modellizzazione della diffusione del calore in una lamina. Tuttavia (cf. per esempio [12]) è ragionevole aspettarsi che tale comportamento possa essere modellizzato con lo stesso grado di fedeltà a un comportamento «reale» dei modelli preda-predatore.

Parleremo degli standard per giudicare teorie, modelli e risultati in un paragrafo successivo.

Metodi.

In questo articolo non è possibile presentare neanche un elenco iniziale dei metodi di ricerca didattica a livello universitario. Come indicazione dell'ampiezza del problema, consideriamo il fatto che l'*Handbook of Qualitative Research in Education* [6] ha quasi 900 pagine! Ci sono dei capitoli in quel volume che contengono estese discussioni sull'etnografia (per esempio, come si può capire la «cultura della classe»?), sull'analisi del discorso (quali regolarità si possono trovare in uno studio accurato di conversazioni?), sul ruolo della cultura nello sviluppo della cognizione, e questioni riguardanti la soggettività e la validità. E questo è solo il lavoro qualitativo — c'è ovviamente anche una lunga tradizione di ricerca quantitativa nelle scienze sociali. Il mio obiettivo è piuttosto quello di individuare un orientamento nei tipi di ricerca che sono portati avanti e di suggerire quali tipi di risultati (e relative limitazioni) possono produrre.

Coloro che si avvicinano alla ricerca didattica tendono a pensare in termini di studi sperimentali standard, che coinvolgono gruppi sperimentali e gruppi di controllo e l'uso di statistiche per determinare se i risultati siano significativi. In realtà, l'uso della statistica in didattica è molto più complesso di quello che si potrebbe pensare.

Per alcuni anni, a partire dalla metà del '900, la ricerca nelle scienze sociali (almeno negli Stati Uniti) è stata dominata dall'esempio dell'agricoltura. Il concetto di base era che se due campi di un particolare raccolto ricevevano lo stesso trattamento, ad eccezione di un elemento variabile, allora la differenza nei raccolti poteva essere attribuita a quella variabile. Si pensava quindi che si potesse fare la stessa cosa con la didattica. Volendo dimostrare che un nuovo metodo di insegnamento per un argomento X era migliore, si sarebbe potuto condurre un esperimento su due gruppi di studenti che studiavano X — a un gruppo si insegnava con il metodo standard e all'altro con il nuovo metodo. Se gli studenti che avevano studiato con il nuovo metodo ottenevano risultati migliori, allora si aveva evidenza della superiorità di quel metodo di insegnamento.

Mettiamo per il momento da parte le questioni sollevate nel paragrafo precedente riguardanti gli obiettivi dell'istruzione e il fatto che il vecchio e il nuovo metodo di insegnamento possano non focalizzarsi sulla stessa cosa. Immaginiamo che si riesca a costruire un test che vada bene per entrambi i metodi, vecchio e nuovo. E supponiamo che gli studenti vengano assegnati a caso al gruppo sperimentale o al gruppo di controllo, in modo da seguire le procedure sperimentali standard. Nonostante questo, ci sarebbero ancora potenzialmente dei seri problemi. Se i due gruppi di studenti avessero insegnanti diversi, ogni differenza nei risultati del test potrebbe essere attribuita a due modi diversi di insegnare. Ma anche con lo stesso insegnante ci possono essere tantissime differenze. Ci potrebbe essere una differenza in energia o coinvolgimento: insegnare «la solita vecchia roba» non è la stessa cosa che sperimentare nuove idee. Oppure gli studenti di un gruppo potrebbero sapere che stanno facendo qualcosa di nuovo e di sperimentale. Solo questo potrebbe già comportare differenze significative. (Esiste una vasta letteratura che mostra che se i soggetti sentono che i cambiamenti sono fatti nel loro interesse allora lavorano di più e meglio — indipendentemente da quali siano i cambiamenti). Oppure gli studenti potrebbero risentire del fatto di essere parte di un esperimento.

Ecco un esempio significativo. Qualche anno fa ho sviluppato dei materiali per lo studio autonomo del «Calculus». Alcuni colleghi di un'altra università hanno accettato di farli usare ai loro studenti. Tranne che per due sezioni, gli studenti che avevano usato questi materiali ebbero risultati migliori di quelli che non li avevano usati. Tuttavia in due sezioni non c'era alcuna differenza significativa. Si venne a sapere che la maggior parte dei docenti aveva introdotto con una presentazione favorevole questi materiali, suggerendo agli studenti che sarebbero stati loro utili. L'insegnante delle sezioni per cui non c'era stata differenza li aveva introdotti dicendo «mi hanno chiesto di darvi questo, non so se vi possano essere utili».

In breve, il metodo sperimentale classico può essere problematico nella ricerca in didattica. Per sottolineare solo due delle difficoltà, raramente gli esperimenti chiamati «doppio cieco» (*double blind*) in senso medico (in cui nè il medico nè i pazienti sanno chi è sottoposto al vero trattamento e chi è sottoposto a un trattamento placebo) sono veramente *ciechi*, e molte variabili sperimentali non possono essere controllate in alcun modo rigoroso. (Questo era il punto centrale dell'esempio descritto nel paragrafo precedente). Di conseguenza, i risultati, sia positivi che negativi, sono difficili da interpretare. Con ciò non si vuole dire che tali studi non siano utili o che i lavori statistici su larga scala non abbiano valore — ce l'hanno sicuramente — ma devono essere portati avanti con molta attenzione e i risultati devono esser interpretati con la stessa attenzione. Lavori statistici di grande valore sono quelli che:

a) producono risultati generali riguardanti una popolazione. Per esempio, Artigue [1] afferma che «più del 40% degli studenti francesi che iniziano l'università dicono che se due numeri A e B sono più vicini di 1/N per ogni intero positivo N, allora non sono necessariamente uguali, ma solo infinitamente vicini»;

b) producono una chiara comparazione di due o più popolazioni. Per esempio, i risultati del Terzo Studio Internazionale su Matematica e Scienze documentano la performance di studenti di varie nazioni in merito a una varietà di contenuti matematici;

c) danno consistenza, nel corso del tempo, a risultati che erano stati scoperti precedentemente con osservazioni su più piccola scala.

Ciò che si può notare è che i metodi di ricerca per la didattica della matematica a livello universitario — direi in tutti i settori della didattica — producono molti risultati, e che la combinazione dei risultati prodotti dai diversi studi nel corso del tempo da consistenza ai risultati.

Commenterò meglio questo punto riportando un esempio tratto dal mio lavoro.

Riguarda il «comportamento metacognitivo» o metacognizione: in particolare l'uso efficace delle proprie risorse (incluso il tempo) nel «problem solving».

Ecco un esempio significativo. Molti anni fa, quando uno degli argomenti standard dell'analisi del primo anno erano le tecniche di integrazione, il seguente esercizio era apparso come primo problema in un test assegnato ad un corso universitario numeroso.

$$\int \frac{x}{x^2 - 9}\, dx \,.$$

Ci si aspettava che gli studenti effettuassero la ovvia sostituzione $u = (x^2 - 9)$ e risolvessero il problema in breve tempo. Circa metà degli studenti fece questo. Tuttavia, circa un quarto della classe, vedendo che il denominatore si poteva fattorizzare, provò a risolvere il problema con la tecnica delle frazioni parziali. E ancora, circa il 10% degli studenti, vedendo che il denominatore era della forma $(x^2 - a^2)$, provò a risolvere il problema effettuando la sostituzione $x = 3 \sin\theta$. Tutti questi metodi portano alla risposta corretta, però il secondo e il terzo sono molto lunghi. Gli studenti che usarono queste due tecniche ebbero dei risultati non buoni nel test, per lo più perché non ebbero tempo per completarlo.

Esempi come questo mi portarono a sviluppare materiali didattici focalizzati sulle decisioni strategiche che si fanno quando si lavora con problemi di integrazione. Questo materiale provocò una differenza nel rendimento degli studenti. Questo mise in luce che le scelte strategiche sono importanti nel «problem solving».

Il problema delle scelte strategiche apparve nuovamente quando, come parte della mia ricerca sul problem solving, esaminai alcuni video di studenti che tentavano di risolvere problemi. Sembrava che molto spesso gli studenti leggessero il testo di un problema, scegliessero rapidamente un metodo di soluzione, e poi perseverassero ciecamente con quel metodo anche quando vedevano che non portava ad alcun risultato. Per far diventare queste osservazioni rigorose, sviluppai un «sistema di codifica» per l'analisi dei video sul «problem solving». Questo schema di analisi conteneva un meccanismo per identificare gli episodi della risoluzione del problema in cui prendere una decisione poteva determinare il successo o l'insuccesso del tentativo. Il sistema di codifica era stato definito in modo tale che anche altri ricercatori potessero utilizzarlo, non solo per analizzare i miei video, ma anche per analizzare i loro. Con questo, i ricercatori potevano vedere come il prendere delle decisioni aiutasse o meno i tentativi degli studenti nel risolvere il problema.

Questi schemi di codifica hanno diversi obiettivi. Primo, avere questi schemi permette che la classificazione dei video sia relativamente oggettiva: se due diversi ricercatori ben preparati producono indipendentemente la stessa codifica per lo stesso video, è ragionevole credere che l'interpretazione sia coerente. Secondo, avere uno strumento analitico di questo tipo permette di evidenziare gli effetti dell'insegnamento del problem solving: confronti video di sessioni di problem solving «prima e dopo» possono rivelare se gli studenti sono diventati risolutori più efficienti o efficaci. Terzo, questo tipo di strumento permette di accumulare dati provenienti da vari studi. In breve, il senso dei risultati in questo caso è: la competenza a livello metacognitivo è un fattore molto produttivo nel «problem solving» (2). Per maggiori dettagli si veda [9].

Come detto prima, i risultati della ricerca in didattica non sono «dimostrati» in senso matematico. In più l'utilizzo dei metodi sperimentali o statistici usati nelle scienze fisiche è spesso difficile, data la complessità di ciò che significa per le condizioni didattiche essere «riproducibili». In didattica si trova un'ampia varietà di metodi di ricerca ed esaminando uno dei primi volumi sulla ricerca didattica a livello universitario, [14], si può avere un'idea di tale varietà. Per lo meno il numero e i tipi di metodi sono aumentati, come si vede nei tre volumi della *Research in Collegiate Mathematics Education.* Si trovano, per esempio, resoconti di dettagliate interviste con gli studenti, confronti tra «calculus» tradizionale e quello della riforma, un esame di «laboratori» di «calculus» e uno studio esteso di come uno studente ha sviluppato comprensione di una macchina fisica e dei grafici relativi a questa. Stanno diventando sempre più comuni gli studi che usano tecniche di osservazione antropologiche e altri metodi qualitativi.

Qual è la validità di questi studi e quanto possiamo contare sui risultati da essi prodotti? È quanto discuteremo qui di seguito.

Standard per giudicare teorie, modelli e risultati.

Esiste una grande varietà di risultati e metodi in didattica della matematica. Un'importante domanda è la seguente: quanto credito di può dare ad un risultato particolare? Che cosa costituisce una ragione valida, che cosa costituisce una «dimostrazione al di là di ogni ragionevole dubbio»?

L'elenco seguente è un insieme di criteri che possono essere utilizzati per valutare modelli e teorie (e più generalmente qualunque lavoro teorico o empirico) in didattica della matematica:

(2) In questo caso (comportamento metacognitivo) molti studi hanno mostrato come prendere delle decisioni efficaci durante il problem solving non sia una «cosa spontanea». Queste capacità possono essere imparate, anche se occorre un insegnamento intensivo. Quando gli studenti acquisiscono tali capacità, la loro prestazione nel «problem solving» migliora.

- Capacità descrittiva
- Capacità esplicativa
- Campo d'azione
- Capacità di previsione
- Rigore e specificità
- Falsificabilità
- Riproducibilità
- Sorgenti multiple di prova (triangolazione).

Descriverò brevemente ognuno di essi.

Capacità descrittiva.

Con il termine capacità descrittiva mi riferisco alla capacità di una teoria di catturare «ciò che conta» in un modo che sembri fedele al fenomeno che si sta descrivendo. Come ha sottolineato Gaea Leinhardt [7], la frase «consideriamo una mucca rotonda» potrebbe essere appropriata quando i fisici vogliono considerare la mucca in termini della sua massa gravitazionale — ma non va bene se si vogliono esplorare le proprietà fisiologiche della mucca! Le teorie della mente, il problem solving, l'insegnamento dovrebbero includere gli aspetti rilevanti e importanti del pensiero, del problem solving e dell'insegnamento rispettivamente. A un livello molto generale, domande da fare sono: manca qualcosa? gli elementi della teoria corrispondono a cose che sembrano ragionevoli? Per esempio, se una sessione di «problem solving», un'intervista o una lezione in classe venissero videoregistrate, una persona che abbia letto l'analisi e guardato il video sarebbe ragionevolmente sorpresa nel constatare oggetti che mancano nell'analisi?

Capacità esplicativa.

Capacità esplicativa significa fornire spiegazioni di come e perché le cose funzionano. Una cosa è dire che delle persone saranno o non saranno in grado di risolvere certi tipi di problemi, o anche descrivere dettagliatamente che cosa fanno; cosa molto diversa è spiegare perché. Per esempio, una cosa è dire che si avranno difficoltà a moltiplicare due numeri di tre cifre mentalmente, ma questo non dà informazioni su come e perché si abbiano difficoltà. La descrizione teorica della memoria di lavoro presentata precedentemente contiene una descrizione delle caselle di memoria, una spiegazione dettagliata del meccanismo di spezzettamento e un'attenta descrizione di come le componenti della memoria interagiscono tra di loro. La spiegazione avviene a livello di meccanismo: dice in termini ragionevolmente precisi quali sono gli oggetti della teoria, come sono legati e perché certe cose saranno possibili e altre no.

Campo d'azione.

Con campo d'azione intendo la varietà dei fenomeni di cui la teoria si può occupare. La teoria delle equazioni non sarebbe molto significativa se si occupasse solo di equazioni lineari. Allo stesso modo la teoria dell'insegnamento non sarebbe molto significativa se si occupasse solo di lezioni frontali!

Capacità di previsione.

Il ruolo della previsione è ovvio: un test per qualunque teoria riguarda la capacità di predire dei risultati prima che questi si verifichino. Nuovamente è bene riferirsi alla teoria dell'evoluzione come modello. Le previsioni in didattica e psicologia non sono sempre del tipo di quelle fatte in fisica.

Qualche volta è possibile fare previsioni precise. Per esempio Brown e Burton [4] hanno studiato i tipi di conoscenze scorrette che gli studenti sviluppano quando imparano l'algoritmo americano standard per la sottrazione in base 10. Essi hanno ipotizzato quali particolari costruzioni mentali gli studenti fanno — l'idea è che gli studenti non solo non riuscivano ad usare l'algoritmo standard, ma sviluppavano una serie di varianti scorrette dell'algoritmo e usavano queste. Brown e Burton hanno sviluppato un semplice test diagnostico con la caratteristica che la sequenza delle risposte sbagliate di uno studente rivelava quale algoritmo errato stava usando. Circa la metà delle volte essi erano in grado di prevedere la risposta sbagliata che gli studenti avrebbero ottenuto con un problema diverso, prima ancora che gli studenti risolvessero il problema!

Questo tipo di previsioni fini sulla base di qualcosa così semplice come un test diagnostico ovviamente sono molto rare. Per esempio, nessuna teoria didattica può prevedere esattamente ciò che farà un insegnante in diverse circostanze; il comportamento umano non è così prevedibile. Tuttavia, una teoria didattica può lavorare in modo analogo alla teoria dell'evoluzione. Può suggerire limitazioni e anche il probabile verificarsi di certi eventi.

[Fare previsioni è uno strumento molto utile nel raffinamento di una teoria. Quando si pensa che qualcosa sia impossibile e invece succede, o quando una teoria afferma ripetutamente che è molto probabile che succeda qualcosa e invece non succede, allora la teoria ha dei problemi! Dunque, fare previsioni è un importante strumento metodologico, nonostante si sappia che fare previsioni precise è impossibile.]

Rigore e specificità.

Costruire una teoria o un modello richiede la determinazione di un insieme di oggetti e di relazioni tra di essi. Questo insieme di oggetti astratti e relazioni dovrebbe corrispondere ad un insieme di oggetti e relazioni nel «mondo reale». Le domande rilevanti sono:

I termini sono ben definiti? Se ne riconoscerebbe qualcuno se lo si vedesse? Nella vita reale? Nel modello? Le relazioni tra gli oggetti sono ben definite? Gli oggetti e le relazioni nel modello quanto corrispondono a quelli che dovrebbero rappresentare? Come detto sopra non ci si può aspettare lo stesso tipo di corrispondenza tra modello e oggetti reali che si ha nel caso di semplici modelli fisici. Costrutti mentali e sociali come le caselle di memoria e il «contratto didattico» (il principio che afferma che insegnanti e studenti hanno delle idee implicite riguardo le regole della loro interazione e che queste modellano il loro comportamento) non sono misurabili come la diffusione del calore in una lamina. Però si potrebbero richiedere dei dettagli, sia per quanto riguarda la definizione degli oggetti sia per quanto riguarda la loro interazione. Le relazioni tra di essi sono ben definite o appaiono come qualcosa «di magico»? Ecco un'analogia. Per quasi tutto il diciottesimo secolo la teoria del flogisto sulla combustione — che affermava che in tutti i materiali infiammabili c'è una sostanza incolore, inodore, insapore, e senza peso, che si libera durante la combustione — era ampiamente accettata. (In seguito il lavoro di Lavoisier sulla combustione rifiutò questa teoria). Seppur con qualche perplessità, la teoria del flogisto spiegava un numero ragionevole di fenomeni. Si sarebbe potuto continuare ad utilizzarla, così come i teorici avrebbero potuto continuare a costruire epicicli su epicicli nella teoria delle orbite circolari ([3]). La teoria avrebbe potuto continuare a produrre risultati utili, abbastanza buoni «per tutti i fini pratici». Questo andrebbe bene per la pratica, ma creerebbe problemi per quanto riguarda la teoria. Come nelle scienze fisiche, i ricercatori in didattica hanno l'obbligo intellettuale di mirare ad una maggiore chiarezza e specificità e di cercare casi limite o controesempi che possano far crollare le idee teoriche.

Ecco due esempi veloci. Primo, nel modello costruito dal mio gruppo di ricerca si rappresentano alcuni aspetti delle conoscenze, degli obiettivi, delle credenze e dei processi di decisione dell'insegnante. Gli scettici (noi inclusi) dovrebbero chiedersi: quanto chiara è la rappresentazione? Dopo aver definito i termini nel modello (cioè dopo che si è specificato che cosa sono conoscenze, obiettivi, credenze di un insegnante) ci sono ancora domande quando spieghiamo che cosa un insegnante potrebbe fare in specifiche circostanze o il modello è sufficientemente ben definito in modo tale che altri possano fare le stesse previsioni? Secondo, la «teoria APOS» descritta in [2] utilizza termini come Azione, Processo, Oggetto e Schema. Sapresti riconoscerli se li incontrassi? Sono ben definiti nel modello? Le loro interazioni e trasformazioni sono ben definite? In entrambi i casi le questioni finali sono: quali sono le stranezze per cui questa teoria

([3]) Questo esempio evidenza un altro importante criterio: la semplicità. Quando una teoria richiede una molteplicità di «punti di riferimento», come epicicli su epicicli, questo è un sintomo che ci deve essere qualcosa di sbagliato.

potrebbe essere dello stesso tipo di quella del flogisto? Coloro che usano la teoria la verificano costantemente per capirlo? Simili domande dovrebbero essere fatte per quanto riguarda tutti i concetti usati nella ricerca didattica, come «contratto didattico», «metacognizione», «concetto immagine» e «ostacoli epistemologici».

Falsificabilità.

Il bisogno della falsificabilità — per non fare affermazioni tautologiche o per fare predizioni la cui accuratezza può essere verificata empiricamente — dovrebbe a questo punto essere chiaro. È una questione parallela a quelle affrontate nei due paragrafi precedenti. Un campo fa progressi (e si guarda dalle tautologie) quando espone le sue idee sul filo del rasoio.

Riproducibilità.

La questione della riproducibilità è anche intimamente legata con quella del rigore e della specificità. Ci sono due tipi di problemi: (1) se si riproducessero le stesse condizioni succederebbe la stessa cosa? (2) altre persone, opportunamente istruite, vedrebbero le stesse cose nei dati? In entrambi i casi per rispondere a queste domande è necessario avere procedure e costrutti ben definiti.

La frase (1) è volutamente vaga perché vuole comprendere un'ampia varietà di casi. Nel caso della memoria a breve termine, l'assunto è che ci troveremo in difficoltà se le attività proposte richiedono l'uso di più di nove caselle di memoria. Nel caso dell'analisi sociologica della classe, l'assunto è che una volta che il contratto didattico è stato inteso, le azioni degli studenti e dell'insegnante saranno conformi a questa (in genere tacita) intesa. Nel caso delle credenze, l'assunto è che gli studenti che hanno determinate credenze agiranno in un certo modo in matematica. Nel caso degli ostacoli epistemologici o della teoria APOS, gli assunti sono similmente che gli studenti che hanno (o non hanno) certe costruzioni mentali saranno (o non saranno) in grado di fare determinate cose.

In tutti questi casi, l'utilità dei risultati, l'accuratezza degli assunti e l'abilità di falsificare o di replicare dipendono dalla specificità con cui i termini sono definiti. Consideriamo un esempio dalla letteratura classica sulla didattica. La teoria di Ausubel degli «advance organizers» (si tratta di quanto è predisposto per l'assimilazione di nuovi metodi e concetti, *n.d.r.*) descritta in [3] sostiene che, se si introduce il materiale che gli studenti devono leggere in modo tale da orientarli in ciò che segue, la loro comprensione del testo migliora in modo significativo. Dopo una decina d'anni e moltissimi studi, la letteratura su questo argomento ancora non conteneva conclusioni: circa la metà degli studi mostrava che essere degli «advance organizers» faceva differenza, l'altra metà diceva invece che non c'era differenza. Un più accurato esame degli studi mostrava il perché: la stessa definizione del termine in questione era piuttosto debole. Diversi sperimentatori hanno prodotto i loro «advance organizers», basandosi su ciò che pensavano avrebbero dovuto essere — e c'erano enormi differenze. Non è sorprendente che i risultati fossero inconcludenti! (Una tecnica standard per ovviare a problemi di buona definizione, e che riguarda la domanda (2) sopra, è di far analizzare i dati indipendentemente a più ricercatori e poi confrontare i risultati. Ci sono delle regole standard per l'affidabilità; queste norme quantificano quanto uguali sono le cose che i ricercatori indipendentemente vedono nei dati.)

Sorgenti multiple di prova (triangolazione).

Qui troviamo una delle più grandi differenze tra la matematica e le scienze sociali. In matematica un argomento convincente (una dimostrazione) è sufficiente: serve a stabilire la validità. Nella didattica e nelle scienze sociali si cerca piuttosto una *prova convincente*. Il fatto è che l'evidenza può essere fuorviante: ciò che pensiamo essere generale potrebbe essere invece frutto di particolari circostanze piuttosto che un fenomeno generale.

Ecco un esempio. Qualche anno fa ho fatto una serie di video con degli studenti del college che lavoravano sul problema: quante cellule ci sono in un corpo umano adulto di taglia media? Il loro comportamento è stato sorprendente. Un certo numero di studenti fece dei tentativi per indovinare l'ordine di grandezza della dimensione di una cellula — da «supponiamo che una cellula sia un angstrom di lato» a «supponiamo che la cellula sia un cubo delle dimensioni di 1/100 di un pollice». In seguito, avendo finito con le dimensioni della cellula abbastanza rapidamente, passarono tanto tempo a lavorare sulla grandezza del corpo, spesso spezzettando il corpo in un insieme di cilindri, coni e sfere e calcolando attentamente il volume di ciascuna di queste. Questo era un procedimento molto strano.

Qualche tempo dopo iniziai a filmare degli studenti che risolvevano problemi in coppia invece che individualmente. Non ho più rivisto il tipo di comportamento descritto sopra. Si scoprì che, quando lavoravano da soli, gli studenti si sentivano sotto pressione. Sapevano che un professore di matematica avrebbe esaminato il loro lavoro. In quelle circostanze, essi si sentivano obbligati a dover fare *qualcosa* di matematico, e calcolando dei volumi almeno sembrava che stessero facendo della matematica! Lavorando a coppie, invece, gli studenti iniziavano dicendo cose come «sicuramente questo è un problema strano». Ciò era sufficiente per allentare la pressione, quindi per loro non c'era più bisogno di calcolare dei volumi per sentirsi sollevati. In breve, un determinato comportamento era funzione delle circostanze e non del problema o degli studenti.

Un modo per controllare i comportamenti artificiali è quello di variare le circostanze e chiedersi: si vede la stessa cosa in tempi diversi e in luoghi diversi? Un altro modo è quello di cercare quante più sorgenti di informazione possibili sul fenomeno in questione e di vedere se queste formano un messaggio coerente. Nel lavoro del mio gruppo di ricerca sulla modellizzazione dell'insegnamento, per esempio, facciamo inferenze sul comportamento dell'insegnante a partire da videoregistrazioni dell'insegnante in azione — però, facciamo anche interviste all'insegnante, esaminiamo il piano delle sue lezioni e il lavoro in classe e discutiamo i nostri risultati con l'insegnante. Più fonti indipendenti di conferma dei risultati ci sono, e più forti diventano i risultati.

Conclusioni.

Il punto principale di questo articolo è che la ricerca in didattica della matematica (a livello universitario) è *molto* diversa dalla ricerca in matematica e che comprendere le differenze è una cosa essenziale se si vuole capire la ricerca in questo campo (o anche contribuire ad essa). I risultati non sono quasi mai definitivi; solitamente sono dei suggerimenti. L'evidenza non è quella dell'ordine di una dimostrazione, bensì è cumulativa, cercando di progredire verso conclusioni che siano valide al di là di ogni ragionevole dubbio. Un approccio scientifico è possibile, però bisogna fare attenzione a non essere degli *scienziati* — ciò che conta non è la bardatura della scienza, come il metodo sperimentale, ma l'uso di un ragionamento attento e di criteri standard per l'evidenza, utilizzando una varietà di metodi appropriati al problema da affrontare.

È utile ricordare che la ricerca in didattica della matematica è un settore molto giovane. I matematici sono abituati a misurare la genealogia della matematica in secoli, o addirittura in millenni; al contrario la genealogia della ricerca in didattica della matematica (in particolare a livello universitario) è misurata in decadi. La rivista *Educational Studies in Mathematics* è datata anni '60. Il primo numero del primo volume del *Journal for Research in Mathematics Education* è stato pubblicato nel gennaio del 1970. La serie di volumi *Research in Collegiate Mathematics Education* — la prima collezione di volumi dedicati solamente alla didattica della matematica a livello di college — iniziò nel 1994. Non è un caso che la maggior parte degli articoli citati da Artigue [1] nella sua analisi dei risultati di ricerca fatta nel 1999, siano stati scritti negli anni '90: c'era ben poco a livello universitario prima di allora! C'è stato un grande progresso negli ultimi anni, però il campo è ancora molto giovane e c'è ancora molta strada da fare.

Data la natura di questo campo di ricerca, è utile ripensare la propria posizione nei confronti del lavoro e della sua utilità. I matematici che si avvicinano a questo lavoro dovrebbero essere aperti a una varietà di idee, capendo che i metodi e le prospettive a cui sono abituati non possono essere trasferite direttamente alla ricerca didattica. Essi non dovrebbero cercare risposte definitive ma idee che possano essere utilizzate. Allo stesso tempo, tutti coloro che usano e fanno ricerca in didattica della matematica (a livello universitario) dovrebbero dare prova di sano scetticismo. In particolare, dato che non ci sono risposte definitive, si dovrebbe diffidare di chiunque ne offra. Più in generale, l'obiettivo principale per le prossime decadi è quello di continuare a costruire un corpo di teorie e metodi che pemettano alla ricerca in didattica della matematica di diventare un campo sempre più robusto, sia come ricerca di base che come ricerca applicata.

BIBLIOGRAFIA

[1] M. ARTIGUE, *The Teaching and Learning of Mathematics at the University Level: Crucial Questions for Contemporary Research in Education*, Notices Amer. Math. Soc., **46** (1999), 1377-1385. M. ARTIGUE, *L'insegnamento e l'apprendimento dell Matematica a livello universitario* (traduzione del precedente lavoro), Bollettino U.M.I., La Matematica nella Società e nella Cultura, Serie VIII, Vol. III-A, Aprile 2000, 81-103.

[2] M. ASIALA - A. BROWN - D. DE VRIES - E. DUBINSKY - D. MATHEWS - K. THOMAS, *A framework for research and curriculum development in undergraduate mathematics education*, Research in Collegiate Mathematics Education (J. Kaput, A. Schoenfeld and E. Dubinsky, eds.), vol. II, Conference Board of the Mathematical Sciences, Washington, DC, pp. 1-32.

[3] D. P. AUSUBEL, *Educational Psychology: A Cognitive View*, Holt-Reinhardt-Winston, New York, 1968.

[4] J. S. BROWN - R. R. BURTON, *Diagnostic models for procedural bugs in basic mathematical skills*, Cognitive Science, **2** (1978), 155-192.

[5] R. G. DOUGLAS (ED.), *Toward a Lean and Lively Calculus*, MAA Notes Number 6, Mathematical Association of America, Washington, DC, 1986.

[6] M. LECOMPTE - W. MILLROY - J. PREISSLE, *Handbook of Qualitative Research in Education*, Academic Press, New York, 1992.

[7] G. LEINHARDT, *On the messiness of overlapping goals in real settings*, Issues in Education, **4** (1998), 125-132.

[8] G. MILLER, *The magic number seven, plus or minus two: some limits on our capacity for processing information*, Psychological Review, **63** (1956), 81-97.

[9] A. H. SCHOENFELD, *Mathematical Problem Solving*, Academic Press, Orlando, FL, 1985.

[10] A. H. SCHOENFELD (ed.) *Student Assessment in Calculus*, MAA Notes Number 43, Mathematical Association of America, Washington, DC, 1997.

[11] A. H. Schoenfeld, *On theory and models: The case of Teaching-in-Context*, *Proceedings of the XX Annual Meeting of the International Group for Psychology and Mathematics Education* (Sarah B. Berenson, ed.), Psychology and Mathematics Education, Raleigh, NC, 1998a.

[12] A. H. Schoenfeld, *Toward a theory of teaching-in-context*, Issues in Education, 4 (1998b), 1-94.

[13] A. H. Schoenfeld, *Models of the Teaching Process*, Journal of Mathematical Behavior (in stampa).

[14] D. Tall (ed.), *Advanced Mathematical Thinking*, Kluwer, Dordrecht, 1991.

Alan H. Schoenfeld, Elizabeth and Edward Conner Professor of Education
Graduate School of Education University of California Berkeley
CA 94720-1670, USA