

Point and Click or Grab and Heft: Comparing the Influence of Physical and Virtual Materials on Elementary School Students' ability to Design Experiments

Lara, M. Triona and David Klahr

小学校におけるコンピュータの普及：アメリカの公立小学校で4人に1台の教育用コンピュータが存在700を超える推奨ソフトウェア（Education Week, 2002）があり，数百が科学教育に関するものである

コンピュータベースのインストラクション：実験科学の教育において明確な利点

携帯性，安全性，費用効果，エラーの最小化，時間と空間の増幅&減少，柔軟性，ダイナミックなデータ表示

コンピュータの普及，利点とは対比的に，コンピュータベースのインストラクションに対する批判

- ・ 実利的，理論的観点からの批判(Armstrong & Casement, 1998, Healy, 1999)
 - よくても効果がないだけで，最悪な場合は有害
- ・ アンチコンピュータの主張
 - Physical な素材を操作することが学習の本質
 - 具体的な操作，実体験の効果（e.g., Berk, 1999, Diem, 2001）

教育において，どのように操作が学習に関係するのかということはまだ始まったばかり
2例の紹介

Chao, Stringer, Woodward(2000)

- ・ 幼稚園児の数の概念学習において，Physical な操作が効果的
しかし，操作，非操作については検討していない

Rensick(1998)

- ・ Physical というよりも，操作が重要
- ・ “digital manipulative”(Virtual な素材)を提唱

Physical vs. Virtual の他の側面

Clark(1983,1994)：1970年代のテレビ，ラジオ，本，初期のコンピュータのメタ分析に基づき，以下の主張

- ・ インストラクションの方法がコントロールされていれば効果に差はない
- ・ これまで差があったのはコントロールに失敗していた

Clark の主張が正しいにせよ，正しくないにせよ，メディアを使用した手法の混乱は残っている．

Virtual vs. Physical 2つの観点

- (a) コンピュータによる素材の提示は、学習を低減させる。なぜなら、Physical なインタラクションが学習の本質であるからである
- (b) 素材が、Physical でも Virtual でも、インストラクションの方法に違いがなければ違いはない

EMPIRICAL STUDIES

直接、コンピュータまたは Physical な素材を使用することの効果を示したものは驚くほど少ない

コンピュータベースの教育システムは多分野において印象的な結果を示しているが

(e.g., Koedinger, Anderson, Hadley, & Mark, 1997, White & Frederiksen, 1998)

これらの評価は典型的、意図的に、コンピュータ対非コンピュータの比較を行ったものである。

上記2つのプロジェクト

- ・ 学習プロセスにおけるコンピュータの役割は不明瞭
他の変化と独立していない

同様の話題

Physical な操作の使用と、Physical な素材を用いない伝統的な教育との比較

メタ分析によると...Physical な操作を用いる効果がある(Bredderman, 1983; Shymansky, Klye, & Alport 1983)

しかし、これらの研究は伝統的な教育との比較で、条件間でカリキュラムが異なっている。

これらの混乱を取り除いた Sowell(1989)による研究

- ・ 絵の操作(e.g., flashcard)と Physical(e.g., 木製ブロック)な操作の比較
- ・ 絵の操作、Physical な操作、伝統的な教育間に差はなし

具体的な操作の評価：数学や科学の領域をカバー

これらの領域における Physical な素材の影響を調べることは可能であるが、そこでは、インストラクションの方法から Physical な操作を取り除くことはできない。

(Virtual が行われていないということ???)

THEORETICAL PERSPECTIVE

Physical な素材の操作...Piaget(1960)の具体的操作期における中心的な役割

構成主義の理論

学習におけるアクティブな役割の重要性。しかし、Physical な操作は必要とされない

- ・ 明白に、目標となるのが知覚・運動スキルであれば操作が効果的であるが、違うタイプのタスクであれば、必要ないであろう。
- ・ 一般的な仮定である、Physical な素材を操作することが学習を促進（特に幼児）するということは、構成主義にも認知学習理論にも根付いていないといえる。

なぜ、Physical な操作がこんなに一般的なのであろう？

可能性

- ・ Physical な素材を用いた方法で効果があったためだろう
伝統的な講義（受身的）に比べ、アクティブであるため
- ・ 具体的な素材の使用
生徒により広範囲な経験を提供

Clements(1999)による提唱

“ digital manipulative を含め、具体的な操作を再定義する必要がある ”

実際に、これまでの研究の多くが、Physical な操作に他の要因が絡んでいる可能性がある。

(Baxter, 1995; Char, 1991; Ferguson & Hegarty, 1995, など...)

これより、Virtual と Physical の効果が同等なのかについての疑問はまだ残る。

THIS EXPERIMENT

本研究：プレゼンテーションの異なる 2 条件の比較

条件 1：Physical で操作できる素材

条件 2：Virtual でコンピュータベースの素材

対象とする領域：実験計画、変数統制方略（以下 C V S）

（実験計画の例：図 1（混乱のない実験）、図 2（混乱した実験））

ほとんどの小学校の科学において、実験計画の授業は存在。しかし...

大多数の小学生は混乱した実験計画を行うことが示されている（e.g., Kuhn, Garcia-Mila, Zohar, & Anderson, 1995; Schauble, 1996）

本研究

Chen & Klahr(1999)のトレーニングの手続きを基礎とする。

Chen & Klahr(1999)の概要

- ・ 小学校 2, 3, 4 年生を対象に C V S を教える
- ・ 素材：バネ, 斜面, 浮き沈み（全て Physical）
- ・ 結果 インストラクション前：25%が混乱のない実験 インストラクション後：約 75%
- ・ 7 ヶ月後の転移テストで、4 年生の学生が非訓練群との優位な差を示す
- ・ C V S パフォーマンスの向上に伴い、領域知識が向上

Toth, Klahr, and Chen(2000)...Chen & Klahr(1999)を通常の科学の授業に拡張

- ・ Chen & Klahr の生徒 / 教師が 1 : 1 から , 20 : 1 になっても同様の効果

これらの研究 : Physical な素材の使用 本研究 : Physical な素材 vs. Virtual な素材

METHOD

Participants

- ・ 南ペンシルバニアの小学校 4 , 5 年生 92 名 平均年齢 10.6 歳 女性 51 名 , 男性 41 名
- ・ Physical , Virtual の群分けはランダム

Design

- ・ 2 (条件 : Physical vs. Virtual material) × 3 (フェーズ : プレテストとトレーニング , ポストテスト , 転移) の要因計画
- ・ フェーズ : 被験者内要因
- ・ 各フェーズにおいて , 簡単な 4 問の比較実験をデザインする

条件について (トレーニング時)

Physical : リアルなバネ , 重りを用いて実験 (図 2)

Virtual : コンピュータスクリーン上で実験 (図 1)

転移課題の時には , 双方とも Physical な素材を新しい領域 (斜面) において使用 .

Materials

Springs.

実際の素材 図 1

4 つの変数が , 2 つのレベルで設定可能 . (バネに対して 3 つ , 重りに対して 1 つ)

バネ

長さ : 長い , 短い

幅 : 広い , 狭い

ワイヤのサイズ : 細い , 太い

重さ

バネにつける重り : 重い , 軽い

目標 : それぞれの実験で , どの変数が重要かを明確にすること

(例 . バネの長さが , どれだけ伸びるのかに変化をもたらすのかどうかを比較しなさい)

Physical 条件 : 実際に手にとって実験

Virtual 条件 : マウスで選択して実験 (実験結果のビデオは Physical 条件の素材を使用し撮影されている)

Ramps.

転移課題に使用

2つの木製斜面，サイズはおよそ24インチ

変数（バネと同一の構造）

斜面の険しさ：とても険しい，すこし険しい

ボールのころがる長さ：長い，短い

斜面の表面：スムーズ，ざらざら

ボール：ゴルフボール，ゴムボール

Confidence scale.

確信の度合いを質問以下の4段階（4点）で聞く

4：完全に確信

3：かなり確信

2：ある程度確信

1：確信をもてない

Procedure

3フェーズ，個人実験

- | | |
|------------------|-----------|
| (a) プレテスト，トレーニング | 45分 |
| (b) ポストテスト | 45分 |
| (c) 転移 | 20分（1週間後） |

Pretest and training

プレテスト

計測するもの...バネに対する初期知識，CVSについての初期知識

手続き

- ・ インストラクターがバネを提示（条件により，Physical or Virtual）し，4つの変数（上記）を紹介
- ・ それぞれの変数について理解度の測定（例．太いまたは，細いときバネを教えてください？）
- ・ 生徒はそれぞれについて確信度を述べる（確信度の練習のため）
- ・ 領域知識についての質問（例．長いバネと，短いバネではどちらがよく伸びますか？）
- ・ 実験のデザインと実行
 - 特定のターゲット変数がバネの伸びに関連するかの実験
 - 4つの実験をデザイン...内2つはバネの長さ，2つは幅について
 - 実験を実行する前にインストラクターにより，以下のような質問を受ける“なぜあなたはこれら2つのバネと重さを選んだのですか”
 - 実行後，4点で確信度を評定し，なぜそう答えたのかを説明

プレテスト終了後

- ・ インストラクターによる約10分間のCVSについての説明
 - 目標は混乱のない実験を計画することであるということと，正事例，負事例の紹介，説明

Posttest

トレーニングフェーズ終了後の即時的な効果の測定

- ・ 4つの実験をデザイン，実行（内2つはワイヤのサイズ，2つは幅について）
- ・ プレテストと同様に，デザインの説明，結果の観察，確信度の評定を行う
- ・ これらの実験終了後，バネの領域知識についての質問

Transfer

トレーニングの1週間後に行われ，全生徒 Physical な斜面を使用．

手続き

- ・ インストラクターによる4つの変数の説明
- ・ 4つの実験をデザイン，実行（内2つは，斜面のなめらかさ，2つは斜面の険しさについて）
- ・ 領域知識（斜面）の測定
- ・ バネと斜面の類似度に関する質問
 - (a) 今週の斜面の問題において，先週のパネの問題を思い出しましたか？
 - (b) 今日の問題と，先週の問題はよく似ていましたか？まったく違いましたか？
 - (c) 先週行った問題は，今週の問題に役に立ちましたか？

RESULTS

Learning and Transfer of CVS

Physical vs. Virtual

2（トレーニング条件）× 3（フェーズ）の分散分析

（トレーニング：Physical，Virtual，； フェーズ：プレ，ポスト，転移）

混乱のない実験の割合について 結果 図3

- ・ フェーズの主効果 $F(2,89) = 65.93, p < .01$
- ・ トレーニングには効果なし $F(1,90) = .01, ns$
- ・ 交互作用なし $F(2,89) = 0.73, ns$
- ・ プレからポスト，転移にかけての優位な向上
- ・ ポストと転移に有意差なし

最も直接的なCVS知識の測定：混乱のない実験を計画するかどうか

トレーニング条件による差はなし．しかし，形式知には差がある可能性
質問に対する回答から形式的なCVS知識を測定

（ターゲットとなる変数のみを変化させ，あとは全部固定しているかどうか）

結果 図4（一名のみ転移フェーズのデータ紛失のため除外）

- ・ フェーズの主効果 $F(2,88) = 62.44, p < .001$
- ・ トレーニングには効果なし $F(1,89) = .71, ns$
- ・ 交互作用なし $F(2,88) = 0.71, ns$
- ・ プレからポスト，転移にかけての優位な向上
- ・ ポストと転移に有意差なし

フェーズの主効果...トレーニングにおける形式知の表現の向上。また、条件による差はなく、両条件の生徒共にCVSを学んだといえる。

ここまでは、全生徒の平均であったが、個人に目を向けてみよう。 図5

定義

good experimenter...ポストテスト、転移テスト合計8実験のデザインにおいて、7以上混乱のない実験デザインをした人

結果

Physical 条件：50% (23/46)

Virtual 条件：52% (24/46)

条件間で差はなし。($\chi^2(1, N = 92) = .04, ns$)

続いて... good experimenter の定義に全変数をコントロールするということを述べたという点を付け加えるとどうなるか？

結果

Physical 条件：28% (13/46)

Virtual 条件：31% (14/45)

条件間で差はなし。($\chi^2(1, N = 91) = .09, ns$ 一名データ紛失で除外)

これより

- ・ 条件によらず、約50%の生徒が8回中7回以上混乱のない実験を計画
- ・ 条件によらず、約25%が明確にCVSについて述べる事が分かる。

good experimenter の数は変わらなかったが、nonlearners(8回中1回以下の混乱のない実験計画)に差はあるのだろうか？

結果：条件間で差はなし。($\chi^2(1, N = 91) = .60, ns$)

ついで、各フェーズに目を向けてみる 表1

各フェーズにおいて秀でているものを定義：4回中3回以上混乱のない実験計画

フィッシャーの直接確率検定：各フェーズの条件間に差はなし

秀でた生徒の割合は条件によらない

ポストテスト、転移では、同等の比率であるがここにはまだ2つの可能性がある

- (a) フェーズ間を通じて同等のパフォーマンスを維持している生徒
- (b) パフォーマンスが改善された生徒と低下した生徒のコンビネーション

これより、ポストテスト 転移において、秀でたものに変化した人を調べると

Physical 条件：ポストテスト 3 人 転移 2 人 差はなし $F(1, N = 46) = .20, ns$

Virtual 条件：4 人改善，4 人降下 差はなし $F(1, N = 46) = .00, ns$

(a)である。

結果を統合すると

- ・全体，個人どちらを見ても，条件による差は見られない

Confidence About Conclusions

これまでの結果から，学習&転移...両群で差はなし

確信度に違いはあるのだろうか？

結果．図 6

2 (トレーニング条件) × 3 (フェーズ) の分散分析

- ・ トレーニングの主効果なし $F(1,90) = 1.21, p = .27$
- ・ フェーズの主効果あり $F(2,89) = 12.68, p < .001$
- ・ 交互作用有意傾向 $F(2,89) = 12.68, p = .06$
- ・ プレテストとポストテスト，転移間に有意差
- ・ ポストテスト，転移間に差はなし

これより

- ・条件間による差はなく，トレーニングにより確信度が上がる。

プレテストからポスト&転移にかけての確信度の向上は，混乱のない実験が増えたから？

混乱のない実験を行った後の確信度の平均を比較

11名は分析対象から除外(3名はすべて混乱がなく，8名はすべて混乱していたため)

- ・ 混乱あり $M = 2.8, SD = .63$
- ・ 混乱なし $M = 3.2, SD = .56$
- ・ 混乱あり，なし間で有意差 $t(80) = 4.86, p < .001$
- ・ 条件間の差はなし $t(79) = 0.38, ns$

結果

- ・ 生徒がより確信を持つようになったのは，条件によらず，混乱のない実験を行ったため

Domain Knowledge

C V S について条件間に差はなかったが、領域知識には差がでるか？

結果・図7

質問内容：4つの変数に対する質問

例．幅のあるばねと細いばねではどちらがより伸びますか？

- ・ 2 (条件) × 2 (時間) の分散分析
- ・ 時間の主効果 ばね： $F(1,90) = 13.99, p < .001$, 斜面 $F(1,90) = 17.44, p < .001$
- ・ 条件の効果なし ばね： $F(1,90) < 0.01$, ns , 斜面 $F(1,90) = 1.46, p = .23$
- ・ ばね交互作用あり $F(1,90) = 6.61, p = .01$
- ・ 斜面交互作用なし $F(1,90) = 1.54, p = .22$

ばねにおいて交互作用...条件による領域知識の学習に差の可能性

しかし、初期($t(90) = 1.69, p = .09$)、最後($t(90) = 1.31, p = .19$)における条件間に差はなく、どのように解釈するのは不明...

全体 個々の変数のパフォーマンスの測定

2つの理由

- (a) 全ての変数において知識を得たのか？特定の変数の知識を得たのか？
- (b) 交互作用は特定の変数に限定されるのか？

²乗検定の結果．8変数の内、4つの変数で有意差あり、より多くの生徒が正確 不正確よりも不正確 正確が多くなった．

		不正確	正確	正確	不正確
・ バネの幅	$p < .001$	60%	7%		
・ バネの重さ	$p < .001$	2%	20%	(悪化)	
・ 斜面のボールの転がる長さ	$p < .001$	22%	4%		
・ 斜面のボールタイプ	$p < .02$	30%	14%		

(他の変数については、初期から80%以上の成績)

これより

バネにおける、時間の主効果は、バネの幅によるものであるが、バネの重さが効果の規模を減少させているといえ、斜面における、時間の主効果は、ボールの転がる長さやボールのタイプによるものといえる．

続いて、その効果は条件によって差があるのだろうか？

²乗検定の結果

両条件において、バネの幅について優位な向上 (不正確 正確)

Physical 条件： $F(1, N = 46) = 24.03, p < .001$ 改善 70% 悪化 7%

Virtual 条件： $F(1, N = 46) = 14.44, p < .001$ 改善 48% 悪化 7%

しかし、ボールのタイプにおいて、異なる効果 $F(1, N = 92) = 8.00, p < .005$

Physical 条件：改善 33% 悪化 7%

Virtual 条件：改善 28% 悪化 22%

斜面においては、全員 Physical な素材を用いたので、なぜこのような差がでたのかは不明...

領域知識のまとめ

- ・ 領域知識の改善は、8変数中の4変数が起因
- ・ 条件間による強固な差はみられない

Similarities Awareness Between Training and Transfer

転移テストの最後に行われた類似性の判定について

気づいた人

Physical : 38% (17/45)

Virtual : 33% (15/45)

$F(1, N = 90) = .19, ns$

類似性についてまとめ

- ・ 両条件とも同程度に気づいていた

DISCUSSION

本研究の結果

小学校4, 5年生において、Physical, Virtual どちらの条件においても

- ・ 混乱のない実験のデザイン
- ・ 実験のデザインと、口頭による正当性の評価
- ・ 転移テスト

において同等の成績を示す

Physical と Virtual の同等の効果

Physical な操作が学習改善という一般的な仮定に対する挑戦

本研究の結果から

Virtual な素材においても、インストラクションの重要な要素を捕らえていれば、Physical なインタラクションは不必要であると言える

教育メディアは学習に関連しないのか？ Clark(1983,1994)の答えは No だった。

しかし、我々はその質問に答えるには2つのことに答える必要があると考える

(a) 異なるタイプのタスクに対するメディアの効果は異なるのか

(b) 人間のインストラクターがいなくてもコンピュータは同様の効果を示すのか

本研究の対象はC V Sであった。

C V Sは領域一般的なスキルであり，小学校の教育において中心的な目標である(National Research Council,1996)

本研究の結果は，C V Sの手続き，概念的知識にはPhysicalなインタラクションが不必要であることを示すが・・・

明白に教育的な目標が知覚・運動能力の獲得であれば，効果は異なり，Physical，Virtualの慎重なコントロールが必要。

人間のインストラクターについて

本研究...両条件共にインストラクターが存在し，コンピュータから直接トレーニングしたときの効果は不明

例．C V Sのステップバイステップのチュートリアル，

教師は，実践型の科学教材を捨てなければならないのか？

そんな必要はないが...本研究で用いたアプローチによって，異なる領域，文脈におけるさらなる調査が必要である。

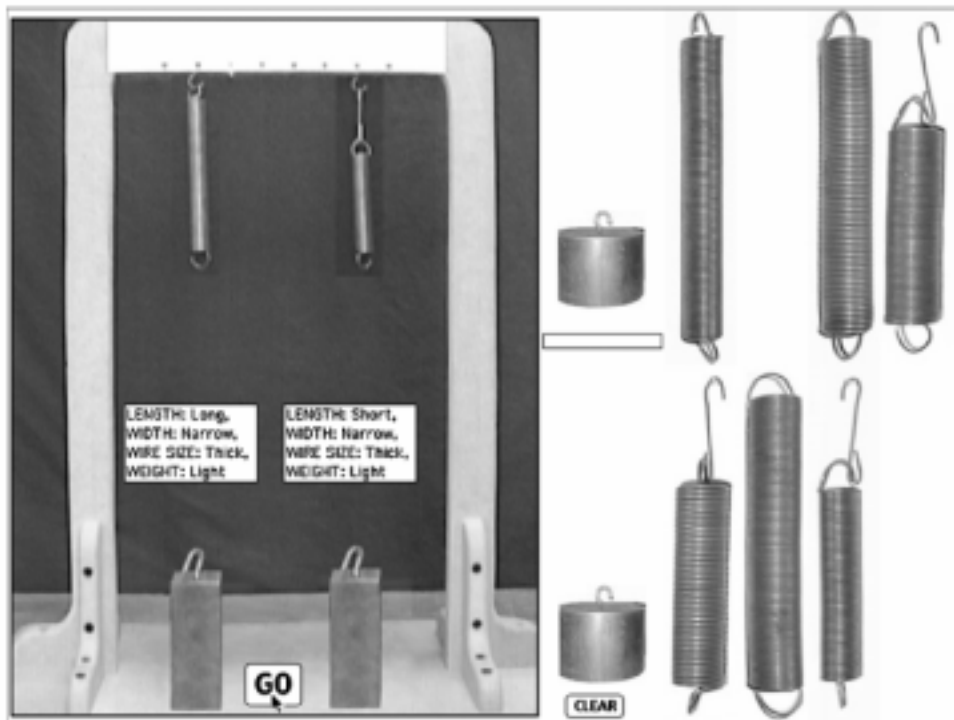


FIGURE 1 Screen shot of the spring selection screen of virtual materials showing an example of an unconfounded experiment of length of spring. Children click on a spring or weight from the right portion of the screen and then click on one side of the hanging rack to select an object for their experiment. After selecting a pair of springs and weights, children click on "go," and the display shows a dynamic video of the springs stretching with the weights.

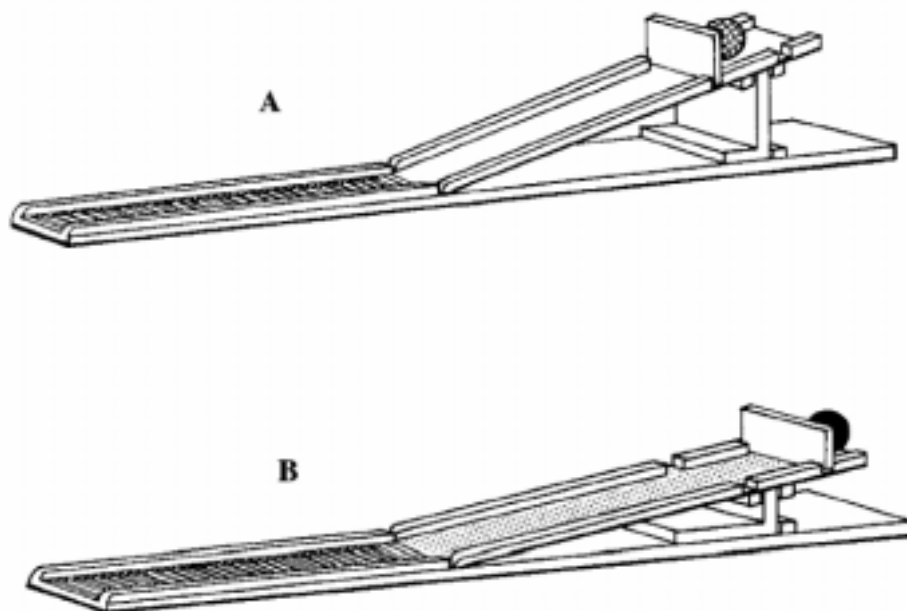


FIGURE 2 Diagram of the ramp materials used during the transfer phase. On each of the two slopes, children can vary the angle of the slope, the surface of the ramp, the length of the ramp, and the type of ball. The confounded experiment depicted here contrasts (A) the golf ball on the steep, smooth, short ramp with (B) the rubber ball on a shallow, rough, long ramp. From "All other things being equal: Acquisition and transfer of the control of variables strategy," by Chen and Klahr, 1999, *Child Development*, 70, p. 1103. Copyright 1999 by Society for Research Development. Reprinted with permission.

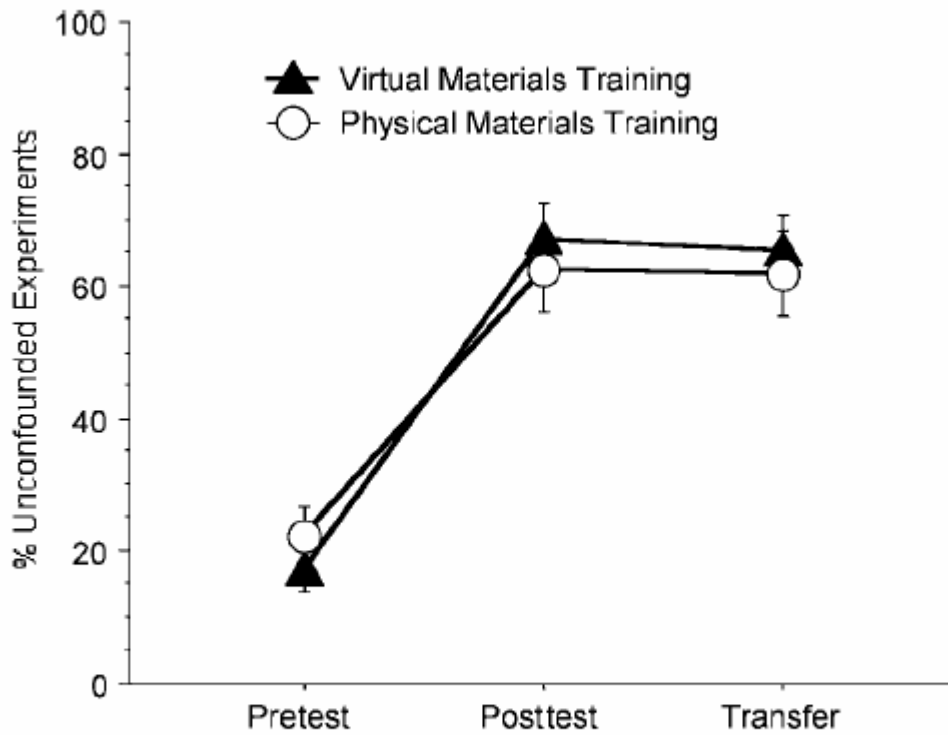


FIGURE 3 Mean proportion of unconfounded experiments for each phase separated by training condition with standard error bars.

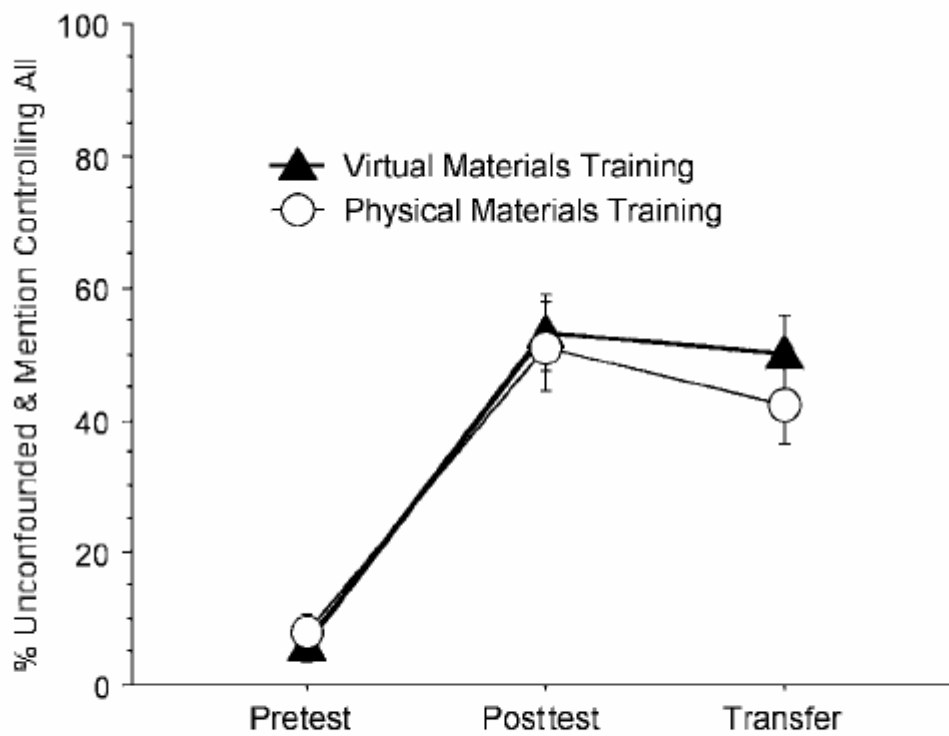


FIGURE 4 Mean proportion of trials in which children designed an unconfounded experiment *and* mentioned controlling all other variables for each phase separated by training condition with standard error bars.

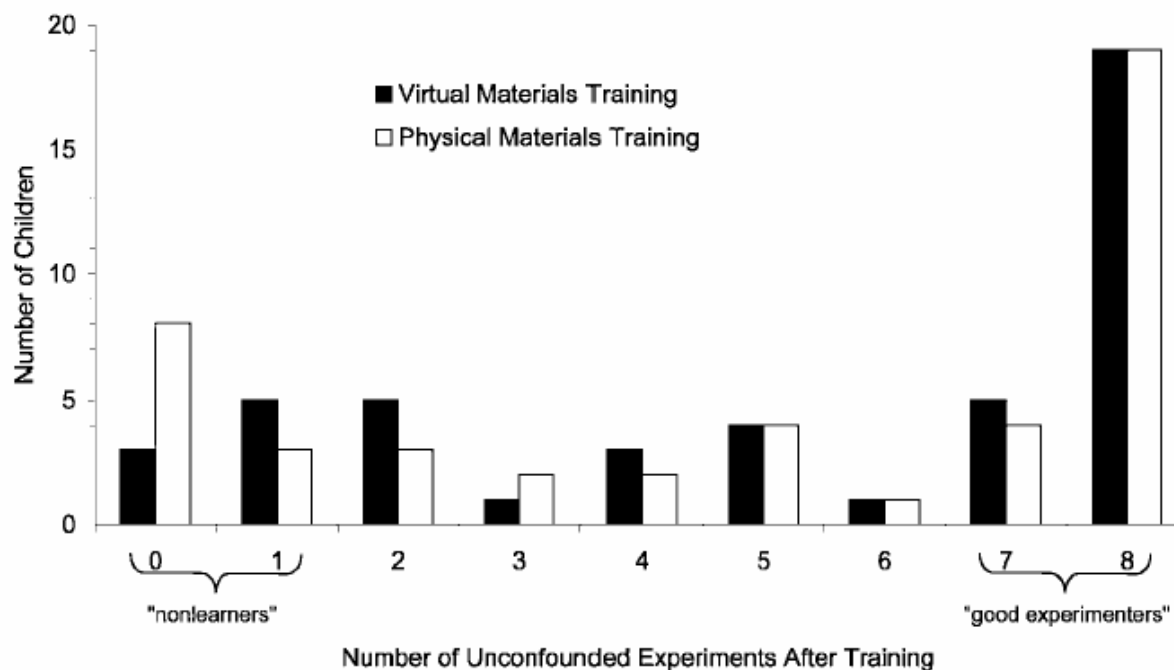


FIGURE 5 Distribution of children separated by condition according to number of unconfounded experiments designed after training (the eight comparisons in the posttest and transfer phases) to show that, although the criteria used to compare the number of good experimenters and nonlearners are arbitrary, the conditions produced similar distributions.

TABLE 1
Proportion and Number of Children With Three out of Four Unconfounded Experiments in Each Phase and Training Condition

Training Type	n	Phase					
		Pretest	%	Posttest	%	Transfer ^a	%
Criterion: Unconfounded experiments on at least three out of four trials							
Physical materials training	46	11	5 _a	57	26 _b	59	27 _b
Virtual materials training	46	4	2 _a	61	28 _b	61	28 _b
Criterion: Unconfounded experiment and mention CVS on at least three out of four trials							
Physical materials training	46	2	1 _c	50	23 _d	35	16 _d
Virtual materials training	45	2	1 _c	52	24 _d	47	21 _d

Note. None of the differences between training conditions are significant using Fisher Exact test $p < .05$. Mean differences in the same row that do not share subscripts differ at $p < .02$ using McNemar chi-square tests. CVS = control of variables strategy.

^aBoth the physical and virtual materials training conditions worked with physical materials during the transfer phase.

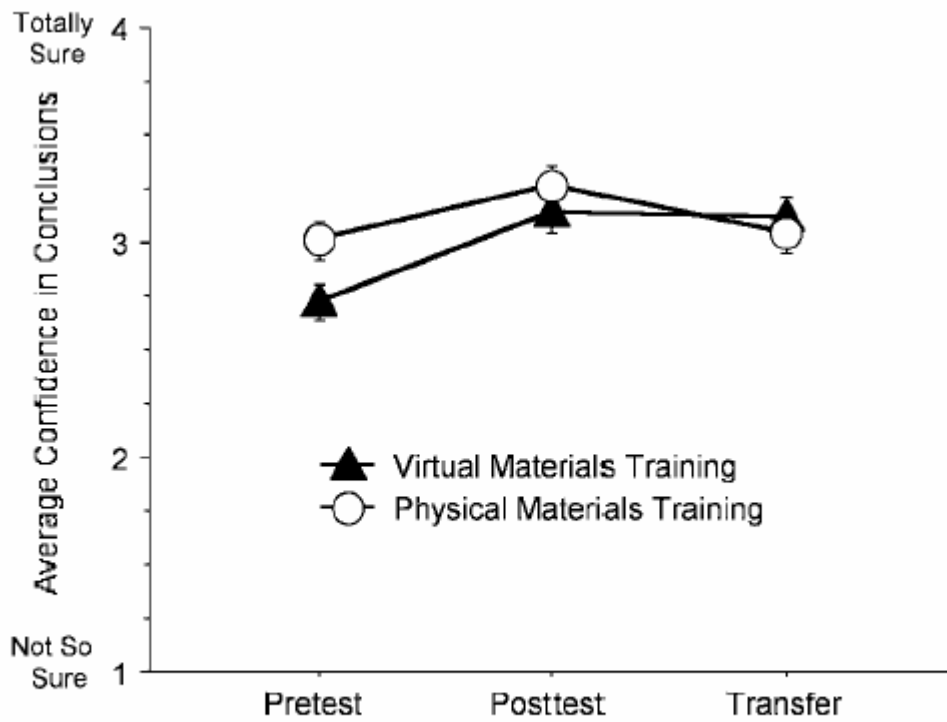


FIGURE 6 Mean of children's confidence in experimental conclusions for each of the phases separated by training condition with standard error bars.

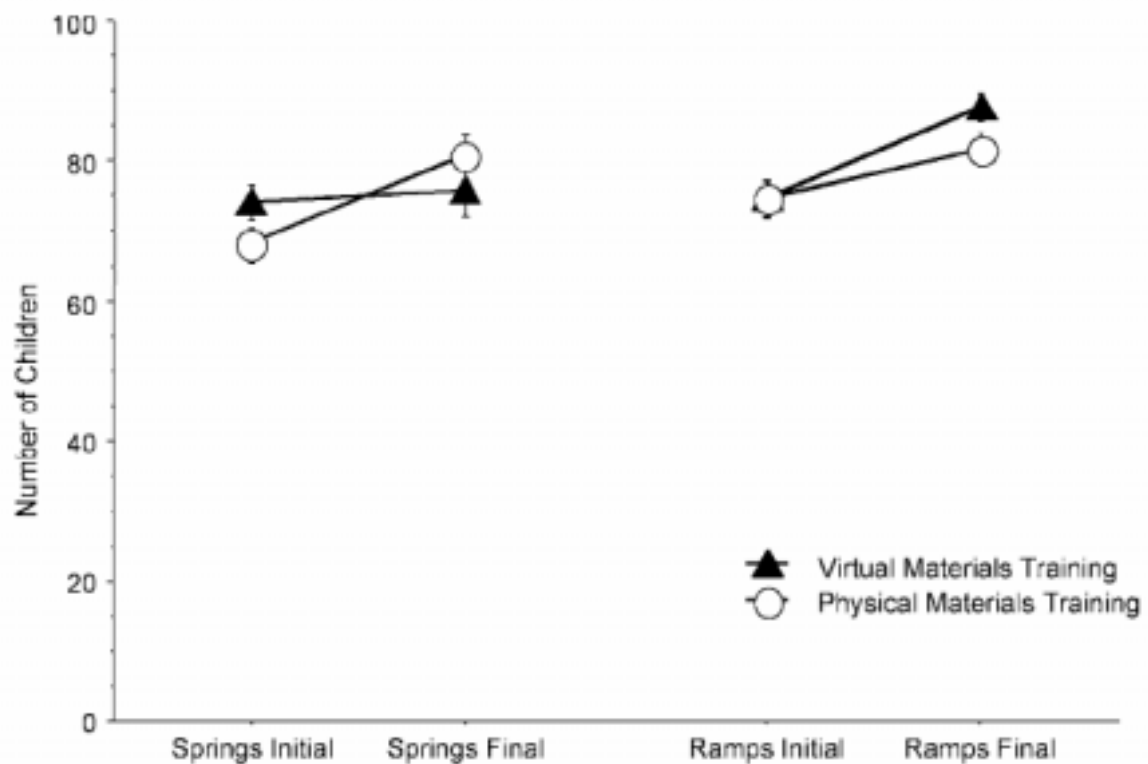


FIGURE 7 Mean proportion of correct descriptions of variable effects for both the training (springs) and transfer (ramps) domains separated by training condition with standard error bars.