

Taking It Out of Context: Collaborating within and across Cultures in Face-to-Face Settings and via Instant Messaging

Leslie D. Setlock, Susan R. Fussell, Christine Neuwirth
Carnegie Mellon University
Proceedings of CSCW 2004 (pp. 604-613)

1 INTRODUCTION

- Computer-mediated collaborations における文化の要因について検討
- 先行研究
 - ◇ メディアを用いた学習や、相互作用に関する研究 (Lundeberg et al, 2000; Teng et al, 1999)
 - ◇ コミュニケーションにおけるリアルタイム性に関する研究 (Massey et al, 2001)
 - ◇ 対面 vs. 非対面状況における文化の影響に関する研究 (Anderson & Hiltz, 2001)
 - ◇ 同一文化 (Homogeneous) vs. 混生文化(Cross-cultural) (Hall, 1976)
- 本研究で取り上げる要因
 - ◇ メディアの要因
 - 対面 (Face to face) vs. コンピュータメディア (Synchronous computer-mediated)
 - ◇ 文化の要因
 - アメリカ人のペア (homogeneous American pairs) vs. 中国人のペア (homogenous mainland Chinese pairs) vs. アメリカン人と中国人のペア (Cross-cultural Chinese/American pairs)

2 DYNAMICS OF COLLABORATIONS

以下では、文化の要因とメディアの要因を検討する際に着目すべき点を取り上げる

2.1 Conversational Grounding

- Grounding
 - ◇ 共通の理解を得るために、様々な事実について言及すること (Clark & Brennan, 1991)
 - ◇ 先行研究では、Common ground (共通の知識, 信念, 目標, 態度, など) を多く提供すればするほど、遂行中の課題の成績が向上することが明らかにされてきている (Clark & Marshall, 1981; Clark & Wilikes-Gibbs, 1981)
- 文化の影響
 - ◇ 文化によって Grounding の仕方が異なる (Triandis, 1980)
 - ◇ 集団主義の文化 (Collective culture)ほど、社会的な事象 (Social cues)に敏感である
- メディアの影響
 - ◇ 対面 vs. コンピュータメディアを比較した研究では、Grounding の難しさを示した研究が数多くある (Hall, 1976)
 - ◇ より Rich なメディアほど Grounding がうまくいくことを示している (Daft, 1984; Strus & McGarth, 1994)

2.2 Message content

- コミュニケーションスタイルの違いに関する検討
 - ◇ 包括的な言葉遣い (Inclusive language) : “We” vs. “I”

- ◇ 丁寧な言葉遣い (Politeness phenomena)
- 文化の影響
 - ◇ 包括的な言葉使い：東洋人は、西欧人に比べて集団主義であるから、“We”を多く使うだろうと考えられる (Morland, 1995)
 - ◇ 丁寧な言葉遣い：文化によって丁寧な単語の使い方が、違う事を示した研究が数多くある (Markus & Kitayama, 1991)
- メディアの影響
 - ◇ Text-based conversations vs. Face-to-face では、コミュニケーションのスタイルに違いが現れる (Keisler et al, 1988)
 - ◇ メディアの種類と文化の種類によって差が表れる (Massey, 2001)
- 2.3 Quality of interaction
 - 相互作用の性質に関する評価
 - ◇ Grounding の仕方や、コミュニケーションのスタイルによって相互作用の評価が変わるのではないか
 - 文化の影響
 - ◇ 文化の異なる者同士のコラボレーションよりも同一文化のコラボレーションのほうが相互作用の性質に対してより良い評価を下すと考えられる
 - メディアの影響
 - ◇ メディアの種類による影響 (Neuwirth et al, 1994)
 - ◇ コミュニケーションを円滑にするメディアほど、相互作用の評価は高くなる (Massey et al, 2001)
- 2.4 Persuasion
 - 説得に関する研究
 - ◇ 自分と似た相手に、説得されやすい (Brock, 1965)
 - 文化の影響
 - ◇ 同じ文化の相手には、説得されやすいのではないか
- 2.5 Task Performance
 - 課題の成績について
 - 文化の影響
 - ◇ ある課題を解いた時の達成感について分析 (Lundeberg, 2000)
 - アメリカ人：実際の結果よりも高く見積もる傾向
 - 中国人：実際の結果よりも低く見積もる傾向
 - メディアの影響
 - ◇ メディアの種類によってパフォーマンスがどのように変わるかを検討した研究 (Fussell et al, 1992)
 - コンピュータを介した状況 vs. 対面では、後者の方が前者よりも成績が良い
- 2.6 The Current Study
 - 研究の目的
 - ◇ 文化の種類 (中国人同士, アメリカ人同士, 中国人とアメリカ人) × メディアの種類 (対面, コンピュータを介した相互作用) を検討する
 - ◇ どのようにすれば, コミュニケーションテクノロジーが異文化コラボレーションを支援できるのか, について模索する
 - 文化の種類と被験者
 - ◇ 中国人とアメリカ人の比較
 - それぞれ文化的に反対の特徴を持つ: 高コンテキスト(中国人) vs. 低コン

- テキスト (アメリカ人)
 - 中国人同士のペア (以後, CC と呼ぶ) とアメリカ人同士のペア (以後, AA と呼ぶ)
- ◇ 同一人種と異なる人種によるペアの形成
 - 文化的背景に着目したペアの形成: 同一人種 (homogeneous) vs. 異なる人種 (heterogeneous)
 - アメリカ人と中国人のペアの形成 (以後, AC と呼ぶ)
- メディアの種類
 - ◇ Instant Messaging
 - コミュニケーションテクノロジーには, Instant Messaging を使用する (以後, IM と呼ぶ)
 - チャットによるコミュニケーション
 - ◇ Face to Face
 - 対面によるコミュニケーション (以後, FF と呼ぶ)
- 予想
 - ◇ Conversational Efficiency (ターンの数)
 - 文化: AA < CC < AC
 - メディア: IM > FF
 - ◇ Conversational Content (We を使う頻度)
 - 文化: CC > AC > AA
 - メディア: IM < FF
 - ◇ Quality of Interaction (アンケートの評価)
 - 文化: 差はないだろう
 - メディア: IM < FF
 - ◇ Persuasion (説得量. ディスカッション前後において意見がどのように変化するのか※詳細は以下に記述)
 - 文化: AA & CC > AC
 - メディア: 差はない
 - ◇ Task Performance
 - 文化: 差はない
 - メディア: IM < FF

3 METHOD

3.1 Design

- 2 要因の混合要因の実験計画
 - ◇ 被験者内要因: メディアの種類 (FF, IM)
 - ◇ 被験者間要因: 人種 (CC, AA, AC)
- 実験状況
 - ◇ 実験室で 2 人の被験者が話し合いをする
 - ◇ 2 種類のメディアを使って課題を行う (課題は 1 ずつ行う. 順番はカウンターバランス.)
 - ◇ 質的, 量的な分析を実施する

3.2 Participants

- 48 人 (Carnegie Mellon University and University of Pittsburgh)
- 米国人と中国人が半数

3.3 Materials

- シナリオ課題

◇ Desert Survival Task/Arctic Survival Task

目的：あるサバイバル状況に関するシナリオが提示される。また、そのときにそのサバイバル状況を克服する上で必要とされる道具が提示される。被験者の目的は、この道具を重要度の高いもの順にランキングをつける事が求められる。

- 全体の手順
 1. トライアル1
 2. トライアル2
 - ※メディアの種類はカウンターバランス
- 各課題での手順
 1. 個別にランキングを行う
 2. 話し合いながらランキングを行う
 3. 個別にランキングを行う
 4. 質問紙
- アイテムの例
 - ◇ Arctic Survival Task : “Gallon can of maple syrup”, “Hand ax”
 - ◇ Desert Survival Task : “Book ‘Edible Plants of the Desert’”, “Loaded .38 caliber pistol”
- 評定尺度
 - ◇ 1 (Most important item for survival in this simulation)から 6 (Least important item for survival in this simulation.)
- 質問紙：コミュニケーションの内容に関する質問
 - ◇ 質問の例
 - “This method of working together was effective”
 - ◇ 評定尺度
 - 1 (Strongly disagree) から 7 (Strongly agree)

3.4 Procedure

- 手順
 - ◇ 実験室への移動
 - ◇ 実験に関する教示（課題，メディア etc）
 - ◇ トライアル1の開始（個人→グループ→個人→質問紙）
 - ◇ トライアル2の開始（個人→グループ→個人→質問紙）

3.5 Coding

- 書き起こされた発言と IM のログを分析
- ターンの分析
- 単語のコーディング
- 分析ツール
 - ◇ Text Analysis and Word Count (TAWC)

3.6 Measures

1. Conversational Efficiency
 - ターンの数のカウント
 - 文字数
2. Conversational Content
 - 会話の内容に関する分析：TAWC を用いて分析
3. Quality of Interaction
 - アンケート：因子分析：要因抽出→3 因子
 1. How well the pair collaborated (e.g., “my partner was responsive to my ideas,”

- “my partner treated me fairly”)
2. How well the pair performed the task (e.g., “we wasted time on this task,” “we agreed on our final answers”)
 3. How frustrated participants felt during the task (e.g., “the task was frustrating”).

4. Persuasion

- グループで話し合う前と後の「ランキング得点」の違い
- 算出方法
 - ◇ 個人で解いた時と、グループで解いた時の差分の絶対値の得点を合計
 - ◇ 2地点でデータを分析：個人→ペア，ペア→個人

5. Task Performance

- ペアでつけたランキングと模範回答 (Human Synergistic) とを比較

4 RESULTS

- 混合要因分散分析
 - ◇ 2(メディア：FF(Face to Face) , IM (Instant Message))×2(人種(以後、条件と呼ぶ) : AA(American-American), AC(American-Chinese), CC(Chinese-Chinese))
 - ◇ なお、課題の種類 (Arctic vs. Desert) については、有意な差が認められなかったため、割愛する

4.1 Efficiency

- ターンと文字数の関係
 - ◇ 相関係数： $r = .89$
 - ◇ 本研究では、ターンにだけ着目する
- 結果：Figure1 参照
 - ◇ 主効果
 - 条件： $AA < AC < CC$, $F [2,19] = 12.41, p < .0001$
 - メディア： $FF > IM$, $F [1,19] = 12.41, p < .005$
 - ◇ 交互作用
 - $F [2, 19] = 5.46, p = .01$
 - 文化による違いは、IM条件において観察されなかった

4.2 Conversation Content

- 結果：
 - ◇ 交互作用：どの項目においても有意差は観察されなかった。以下では、カテゴリーごとに主効果の結果を見ていく
 - ◇ 条件の主効果：Table 1
 - "We" : $CC > AA \ \& \ AC$, $F [2, 19] = 3.64, p < .05$
 - "Social language" : $CC > AA \ \& \ AC$, $F [2, 19] = 6.74, p < .01$
 - ◇ メディアの主効果：Table 2
 - "Pronouns" : $FF > IM$, $F [1, 19] = 15.51, p < .001$
 - "I" : $FF > IM$, $F [1, 19] = 8.88, p < .01$
 - "You" : $FF > IM$, $F [1, 19] = 17.88, p < .0001$
 - "Affective terms" : $FF > IM$, $F [1, 19] = 3.78, p = .07$
 - "Hedges" (意味：言質(げんち)を取られないようなあいまいな発言；緩衝的表現。) : $FF > IM$, $F [2, 19] = 13.06, p < .005$

4.3 Quality of Interaction

- 質問紙の分析：因子項目ごとに分析
- 結果：Table 3
 - Collaboration scale
 - ◇ 主効果
 - 条件：CC > AC & AA, $F [2, 21] = 4.11, p < .05$
 - メディア：[n.s.]
 - ◇ 交互作用
 - [n.s.]
 - Task Performance
 - ◇ 主効果
 - 条件：[n.s.]
 - メディア：[n.s.]
 - ◇ 交互作用
 - [n.s.]
 - Frustration
 - ◇ 主効果
 - 条件：[n.s.]
 - メディア：[n.s.]
 - ◇ 交互作用
 - [n.s.]

4.4 Persuasion

- 算出方法は、上記記載
- 結果：Figure2 ※メディアの要因はつぶしている
- Pre-discussion rankings
 - ◇ 主効果
 - 条件：[n.s.]
 - メディア：[n.s.]
 - ◇ 交互作用
 - [n.s.]
- Post-discussion rankings
 - ◇ 主効果
 - 条件：AC > AA > CC, $F [2, 21] = 3.60, p = .05$
 - メディア：[n.s.]
 - ◇ 交互作用
 - [n.s.]

CCのポストランキングの結果より、グループで合意した内容を個人のラインキングにおいて反映されていることを示す

4.5 Task Performance

- 算出方法は、上記記載
- 結果：
 - ◇ 主効果
 - 条件：[n.s.]
 - メディア：[n.s.]
 - ◇ 交互作用
 - [n.s.]
- 相関分析

- ◇ 得点 × Words : ($r = .26, p < .10$)
- ◇ 得点 × Speaking turns : ($r = -.25, p < .10$)
- ◇ 得点 × “We” pronouns : ($r = .42, p < .005$)
- ◇ 得点 × “I” pronouns : ($r = -.37, p = .01$)
- ◇ 得点 × Discussion of cognitive mechanisms : ($r = -.33, p < .05$)

4.6 Qualitative Analysis of Dialogues

3つの観点から質的に分析を行った。

4.6.1 Depth of analysis

- 「どこまで掘り下げて話し合っていたか」という観点で文化的に違いが観察された：Table4
 - ◇ AA: 順位に関する理由よりも順位そのものに関する議論が中心 (AA pairs were more likely to negotiate placement order rather than analyze the reasons for the ordering.)
 - ◇ CC: 使用用途や、詳細な事柄に関する議論 (CC pairs often discussed the potential uses and value of the items in detail.)

4.6.2 Task investment

- 「選択の正当性」という観点で文化的な差が見られた
 - ◇ AA: できるだけ速く解くということを優先させた (AA pairs sought to reach a negotiated solution to which they could both quickly agree) (以下参照)
P1: ah, wait, I thought it was axe then matches then canvas
P2: okay
P2: that'll work too.

P2: axe is 3?
P2: u want to change to 4
P1: 4
P1: yeah
P2: ok
P2: done [AA20]
 - ◇ CC: Table 4 にあるように合意を形成するまで話し合う傾向がある. (CC pairs sought internal consensus)

4.6.3 Relationship investment

- Table1 にあるように、中国人は「We」をよく使う傾向にある
- →対人関係を構築するような行動をとるのではないか (Relationship-building aspect)
P2: So the first two things are water and pistol.
P1: maybe ☺
P1: totally agree
P2: OK, we have made the first choice! Congratulations!
P2: Then we need to think more
- 例にあるように
 - ◇ 自分たちの状態を確認するようなメタ的な発言をしていた (meta-statements)
 - ◇ →AA や AC では、このような方略はほとんどみられなかった

5 DISCUSSION

- 5.1 Conversational grounding
 - 量的分析：ターンテイキング，文字数
 - ◇ AA のペアは，メディアの種類に関わりなく，より速く Grounding が起きていた
 - ◇ CC のペアは，より時間を要した
 - 質的分析：
 - ◇ CC では，より深い認知的な合意形成が行われていた
 - ◇ Grounding の仕方が異なる：Clark (1986)の見解と一致
- 5.2 Message content
 - 丁寧さ：Hedges に着目
 - 量的分析
 - ◇ 文化的な差は認められなかった
 - ◇ コーディングの方法を変えるべき？
 - 質的分析
 - ◇ CC ペア：相手の意見を聞いたり，意見の不一致を回避しようとする
- 5.3 Quality of interaction
 - アンケートの結果
 - ◇ CC > AC & AA
 - ◇ CC では，議論が深まった事によって，満足のいく相互作用に関する評価を行ったのではないかと考えられる
- 5.4 Persuasion
 - ポストにおける変化量：AC > AA > CC
 - ◇ 深い議論をしていくなかで，意見の不一致に関する理由を納得のいくまで話し合ったからではないか
- 5.5 Task Performance
 - メディア，条件（文化）において何処にも差がみられなかった
 - アイテムがサバイバルという状況において，「どのように役立つのか」に関する知識が乏しかった可能性がある。
 - 今後は，発話から被験者が評価を下せるような知識を持っていたかどうかを再コーディングする予定である。
- 5.6 Limitations and Future Directions
 1. 文化のどの側面が相互作用に影響しているのかを詳細に検討することである。(Autonomy vs. Agreement)
 2. Conversational Efficiency の交互作用より，IM における文化の影響は少なくなることがわかった。ただし，何が影響したのかについて，今後より詳細に検討しなくてはいけない。
 3. 課題に依存する結果。他の種類のコラボレーションでは，より Riche なメディアが必要となるだろう
- 6 CONCLUSION
 - 本研究のいくつかの分析より，CC と AA とでは，シナリオを異なるアプローチで取り組んでいた事が明らかになった
 - ◇ AA：とにかくランキング表を完成させる方略 (exercise in situation-specific compromise)
 - ◇ CC：とことん合意するまで話し，ランキング表を形成する (consensus reaching task)

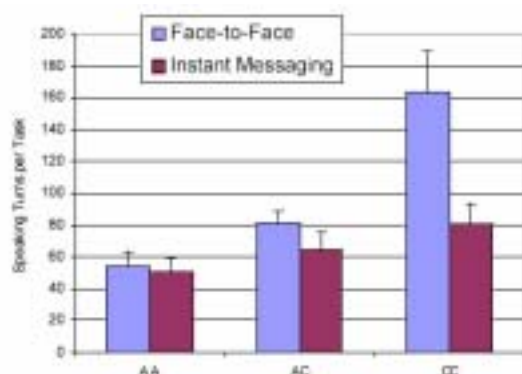


Figure 1: Mean number of speaking turns per trial and pair by culture group and media condition. (AA = American only, AC = Mixed American Chinese, CC = Chinese only.)

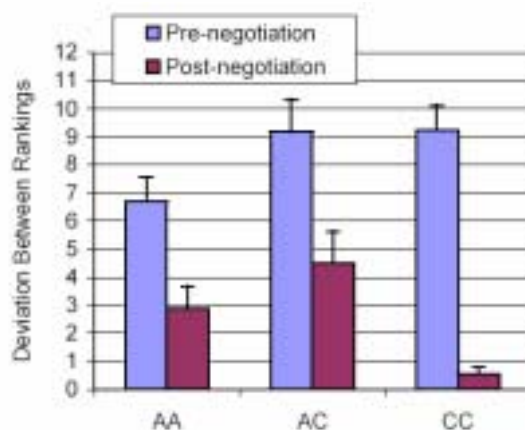


Figure 2: Agreement in pre- and post-discussion private rankings by culture group (AA = American only, AC = Mixed American Chinese, CC = Chinese only).

Table 1. Percentages of words per category by culture group (s.d. in parentheses). Note: Means marked with different superscripts are significantly different from one another.

	AA	AC	CC
Pronouns	11.31 ^a (2.12)	11.43 ^b (2.40)	12.03 ^a (2.56)
I	4.60 ^a (1.58)	3.92 ^a (1.76)	4.40 ^a (0.96)
We	0.86 ^a (0.53)	1.56 ^{a,b} (1.19)	1.99 ^b (0.93)
You	2.32 ^a (1.43)	2.78 ^a (1.62)	3.30 ^a (2.00)
Affective	4.62 ^a (1.34)	4.23 ^a (1.34)	4.30 ^a (1.18)
Cognitive	8.67 ^a (2.32)	8.81 ^a (1.78)	8.96 ^a (1.70)
Social	5.46 ^a (1.46)	5.94 ^a (1.14)	6.95 ^b (1.36)
Hedges	1.29 ^a (1.25)	1.44 ^a (1.28)	1.52 ^a (0.88)

Table 2. Percentages of words by category and media condition (s.d. in parentheses). Note: Means marked with different superscripts are significantly different from one another. [note: the affective differ at $p < .07$]

	FF	IM
Pronouns	12.84 ^a (2.13)	10.34 ^b (1.81)
I	4.96 ^a (1.58)	3.67 ^b (1.12)
We	1.45 ^a (0.98)	2.07 ^a (1.07)
You	3.38 ^a (1.44)	2.21 ^b (0.98)
Affective	3.97 ^a (1.14)	4.17 ^a (1.30)
Cognitive	8.69 ^a (1.66)	8.68 ^a (2.19)
Social	6.34 ^a (1.40)	6.17 ^a (1.44)
Hedges	1.85 ^a (1.00)	0.99 ^b (0.53)

Table 3. Scores on collaboration, task performance, and frustration factors by culture group (1 = lowest, 7 = highest). Note: Means marked with different superscripts are significantly different from one another.

Scale	AA	AC	CC
Collaboration	6.24 ^{ab}	5.81 ^a	6.53 ^b
Task Performance	4.98 ^a	4.42 ^a	5.52 ^a
Frustration	2.13 ^a	2.31 ^a	2.25 ^a

Table 4. Illustrative examples of discussion of the value of the mirror in the desert task from the IM condition. (AA = American only, CC = Chinese only).

AA	CC
P1: why did you put the mirror as 4? P2: you could use it to signal for help or something like that P1: that's true...okay maybe make it 5 then? [16]	P1: what's the use of mirror? P2: so jacket is not last, or mirror is? P2: mirror is for asking help mainly, I guess P1: how about using pistol to ask for help? P1: mirror can be last if we have pistol in front of it. P2: yes, so either mirror or jacket is the least important P1: shoot the sky is better than a mirror. P1: right, P2: ok, mirror P1: I guess jacket is more important than mirror.
P1: mirror last? P2: yeah [12]	P2: and then mirror P2: since mirror helps others to find them P1: yeah

概要

近年、コミュニケーションの技術の発展に伴い、国際的な水準でのコラボレーションが可能となってきた。本研究では、様々な文化によって構成されたペアが2種類のメディアを使って課題に取り組む状況を設定し、実験を実施した。実験には、Carnegie Mellon University と University of Pittsburgh の大学生 48 人（米国人と中国人が半数）が参加した。

実験のデザインは、2 要因混合計画である。被験者間要因として、ペア構成員の文化を操作した。ここでは、次の3条件を設定した。中国人同士のペアと、アメリカ人同士のペア、アメリカ人と中国人のペアである。被験者内要因として、メディアの種類を操作した。ここでは、次の2条件を設定した。Instant Message チャットによるコミュニケーションと、対面によるコミュニケーションである。

被験者は、2 種類のシナリオ課題を行った。この課題では、被験者は、あるサバイバル状況に関するシナリオと、その状況を克服する上で必要とされる道具が提示される。被験者の目的は、この道具を重要度の高いもの順にランキングをつける事である。被験者は、単独でのランキングとペアでの話し合いに基づく2種類のランキングを行った。従属変数として、次の項目が分析の対象となった。1. Conversational Efficiency (ターンの数, 文字数) 2. Conversational Content (会話の内容に関する分析) 3. Quality of Interaction (アンケート) 4. Persuasion (グループで話し合う前と後の「ランキング得点」の違い) 5. Task Performance (ペアでつけたランキングと模範回答とを比較)

質的、量的分析を実施した結果、中国人同士とアメリカ人同士のペアでは、シナリオを異なるアプローチで取り組んでいた事が明らかになった。具体的には、アメリカ人同士のペアでは、とにかくランキング表を完成させる方略 (Exercise in situation-specific compromise) で取り組んでいたのに対して、中国人同士では、とことん合意するまで話し、ランキング表を形成する (Consensus reaching task) という方略で課題に取り組んでいた。