

小学生のコミュニケーション能力に対する Performance Assessment

－ 活動プログラム (Task) と評価基準 (Rubric) 作成の試み －

廣岡 秀一¹⁾・中西 良文¹⁾・廣岡 雅子²⁾・横矢 祥代³⁾
福田 真知⁴⁾・秋山 美和⁴⁾・伊藤 由恵⁴⁾・小倉 明子⁴⁾

本研究では、子どものコミュニケーション能力を評価するために、Performance Assessment (PA) に基づく新たな方法の開発と検討を行った。まず、評価のターゲットとなるスキルが発揮されるパフォーマンス課題である Task とそこで見られる行動の評価基準である Rubric を作成した。そして、著者らに関わる「わくわくコミュニケーションクラブ」の活動の中で Task を実施し、そこでの行動を Rubric を参照しながら評定した。その結果、2回の PA の実施の中で、Rubric・Task とともに改良や工夫が施され、また評定者の訓練効果により、2回目の PA では評定者間の一致率が上昇した。また、Task 中に最も多く見られた行動を評定する最頻行動評定と、Task 中に見られた最高の水準の行動を評定する最高行動評定との間の関連はスキルによって異なり、また他の評定指標である行動チェックリストとの関連もスキルによって異なっていた。これらの解釈については、Task の特徴によるとも考えられるため、今後のさらなる検討を待たなければならないが、コミュニケーション能力評価の新たな方法として、PA に基づく今回の方法に一定の可能性があることが示唆された。

キーワード：コミュニケーション能力、社会的クリティカルシンキング、社会的スキル、実践的プログラムの開発、Performance Assessment

問題と目的

近年、子どもの対人関係能力や対社会関係能力の低下について多くの指摘がなされている (e. g., 吉田, 1997)。こうした認識のもと、われわれ三重大学教育学部教育心理学教室の教員・大学院生・大学生 (以下、スタッフと呼ぶ) を中心メンバーとする「わくわくコミュニケーションクラブ」と称されたボランティアグループは、三重県津市立 M 小学校区内の子どもを対象として、2 年間にわたり、心理学をベースとした小学生のコミュニケーション能力の育成のためのプログラム開発、実践および報告を行ってきた (廣岡・中西・廣岡・後藤・横矢・矢神・福田, 2005a ; 2005b ; 廣岡・中西・廣岡・横矢・福田・秋山・伊藤, 2006)。

この実践では、グループ活動を中心にしながら、人間関係の形成や発展を促したり、自他の経験する感情やその表現に注目させたり、社会的クリティカルシンキング (廣岡・小川・元吉, 2000 ; 廣岡・元吉・小川・斎藤, 2001 など) や社会的スキル (Social Skill) を獲得することにつながるグループ活動を通じて、子どものコミュニケーション能力を高めるためのプログラムの開発と実践を行ってきた。この活動におけるスタッフは、構成的グループエンカウンターリーダー的な役割を担いなが

ら子どもと温かい関係性を築く一方で、社会的な存在としての人間に対する子どもの気づきや発見、自発的な理解を尊重する支援を行ってきた。

さて、このようなコミュニケーション能力を育成することを目的とした活動においては、活動がもつ効果に対する評価がきわめて重要になってくる。心理学的知見に基づいた実践プログラムの開発とその効果の測定の両方に取り組むことで、プログラムの効果だけでなく、評価基準が明確になることから活動の目標がより明確になり、さらに高い効果を生むプログラムの開発につながるというブラッシュアップサイクルの実現が可能となると考えられる。

さて、実際にコミュニケーション能力が育成されれば、子どもの社会的行動に変化が見られるはずである。わくわくコミュニケーションクラブでは、何らかの形で他の子どもやスタッフとのコミュニケーションを行うような体験的な活動を行っている。そのため、活動中の子どもの行動を捉えることにより、子どもが身につけたコミュニケーション能力の度合いやその質を評価することができると考えられる。そこで、わくわくコミュニケーションクラブが子どもにもたらした効果について検討するために、子どもの行動の観察評定がスタッフによって試行的に開始された (廣岡ら, 2006)。

1) 三重大学教育学部 e-mail:shuhiro@edu.mie-u.ac.jp
2) 三重大学教育学部附属中学校非常勤カウンセラー
3) 三重大学高等教育創造開発センター
4) 三重大学大学院教育学研究科学校教育専攻

ところで、教育評価における近年の議論においては、伝統的なペーパーテストによる教育評価がもついくつかの問題を再認識するとともに、その解決策の一つとして、学習者の遂行行動を直接評価しようとするパフォーマンス・アセスメント (Performance Assessment、以下 PA) が注目され始めている。PA とは、ある特定の文脈のもとで、さまざまな知識や技能などを用いながら行われる、その人自身の作品やふるまい (パフォーマンス) を直接に評価する方法のことである (鈴木, 2004; 松下, 2005)。例えば松下 (2005) は、PA の考え方に基づいた Rubric (評価基準) と Task (パフォーマンス課題) を用いて児童、生徒の算数と数学の学力 (課題に対する思考プロセスなど) を測定しようと試みている。PA においては、できるだけ現実的で本物らしい場面を設定し (文脈性)、能力をひとまとまりのものとして把握しようとする (分割不可能性) ことにその特徴がある (鈴木, 2004; 松下, 2005)。コミュニケーション能力についても、現実の社会的場面において発揮されるものであり、また、ある特定のコミュニケーション能力だけを抜き出して検討することが難しいものである。そこで、上述の特徴を持つ PA という方法を適用することによって、コミュニケーション能力の測定が可能になると考えられる。

したがって、本研究は、わくわくコミュニケーションクラブにおける活動効果を測定するために、PA の考えに基づいた Rubric を開発することと、その Rubric を用いて評定できるような Task (パフォーマンス課題となるべき活動プログラム) を開発することを目的とする。

方 法

1. Rubric 作成

- ①ターゲットスキル: 6つのスキルを取り上げた。これらはこの活動が育成を目指すコミュニケーション能力を構成する根幹となるスキルと考えている。すなわち、話すスキル (話し合うとき、自分の考えや意見を相手に伝えようとしているかどうか)、聞くスキル (人の話を聞くときに積極的傾聴ができていないかどうか)、協力スキル (特に作業する場面でグループで協力することができるかどうか)、調整スキル (友だちと意見があわない場合、折り合いをつけて納得できるような提案ができるかどうか)、応答性スキル (相手からの働きかけに適した反応ができるかどうか)、頼むスキル (相手に頼み事をしたいときに、相手に不快感を与えず、自分のしてほしいことをはっきりと伝えながら頼むことができるかどうか) である。
- ②手続き: 松下 (2005) を参考に、各スキルについて、5水準からなる Rubric を Table1 のように作成した。話すスキルと聞くスキルは廣岡ら (2005b, 2006)

を、頼むスキルは吉田・廣岡・齋藤 (2002) や廣岡ら (2005a) を参考にした。協力スキルと調整スキル、応答性スキルはコミュニケーション能力の中心的スキルとして担当者が選択し、評定水準を独自に考案した。数回の検討を経て、スキルの各水準を文章で記述し、具体例や評定のポイントも併記して完成させた。具体例や評定のポイントの例として応答性スキルを Table 2 に示す。

2. 評定される Task の概要

- ①実施日: 2005年9月から同12月まで6回実施した「わくわくコミュニケーションクラブ秋クラス」のうち、第2回活動 (10月1日) と第6回活動 (12月3日) の、「グループによる話し合いと作業」を評定のための Task とした。以下それぞれを Task1、Task2 と表記する。なお、秋クラスの第1回活動では「あいさつのしかた」、第3回活動では「頼み方」、第4回活動は「断り方」、第5回活動では「相手の気持ちを考える」を学習する活動を行った。
- ②実施場所: 津市立 M 小学校図書室
- ③評定対象児: M 小学校の4-5年生11名 (4年男子3名、同女子3名、5年男子1名、同女子4名)。実際の活動は、同じ教室内にて6名と5名の2グループに分かれて行われた。
- ④活動立案・実施者 (スタッフ): 筆者らの他に、三重大学教育学部学生7名、兵庫教育大学大学院学校教育研究科学生1名の、計16名が活動立案・実施者となった。
- ⑤評定されたスキル
Task1: 話すスキル、聞くスキル、協力スキル、調整スキル、応答性スキル
Task2: 話すスキル、聞くスキル、協力スキル、調整スキル、応答性スキル、頼むスキル
- ⑥Task の内容と記録
Task1: グループごとに模造紙にすごろくを作らせた。話し合い場面で話すスキル、聞くスキル、協力スキル、調整スキル、応答性スキルを観察した。具体的には、すごろくのタイトル、マスに書き込む2つの自由お題、スタート・ゴール・マスの配置をグループで話し合い、グループで決定させることを課した。また、実際にすごろくを作成する作業や背景を自由に書き加えさせることによって協力スキルなどをさらに観察する機会を増やした。記録は、グループに1台のビデオカメラによる録画と IC レコーダーによる音声録音によって行われた。評定は、この録画画像と音声を対象に行われた。Task2: 町の地図を作らせた。グループの成員が順に受け取る情報カードをもとに、複数の建物の位置をグループで推理させて不完全な地図を埋め、余白を自由に記入して架空の町の地図を完成させ、答え合わせを

Table 1 ターゲットスキルの Rubric

	Not yet 全く見られない 1	Novice 初心者程度 2	Apprentice まああ 3	Proficient だいたいできている 4	Distinguished 完璧 5
話す	自分の意見を言うという行動は全く見られない。	何らかの意見は言うことができているが、みんなには伝わっていない。適切な大きさの声でない。	聞こえやすい声（適切な大きさの声）で意見を言うことができている。	聞こえやすい声（適切な大きさの声）で、はっきりと最後まで自分の意見を言うことができている。	聞こえやすい声（適切な大きさの声）ではっきりと最後まで自分の意見を言いつつ、目配りができている。
聞く	人の話を聞く姿勢ができていない。姿勢以前に、聞いていない。（聞いているとは判断できない）	聞いているようだと判断できるが、よそ見よそ事をしている。 ※よそ見よそ事；活動とは関係のないこと、文脈とは関係のないこと	人が話をしているときによそ見よそ事をせず、かつ話す人の方を見ている。	人が話をしているときによそ見よそ事をせず、かつ話す人の方を見て、話を最後まで聞いている。	人が話をしているときによそ見よそ事をせず、かつ話す人の方を見て、うなずいたりあいつちをうつなど話に反応しながら、話を最後まで聞いている。（質問も含む）
協力	人の様子にお構いなしに、ひとりで行動している。	グループ全体ですべきことに参加している。	友だち（スタッフ）と一緒に何かをしようとしている。	グループ全体の状態を把握して、意見を言ったり作業したりしている。	グループ全体の状態を把握し、みんなに声をかけながら、手際よく進めることができる。
調整	人の意見を無視し、自分勝手に行動している。	相手と自分の意見が合わなければ黙ってしまう。沈黙。何もしない。	相手と自分の意見が合わなければどちらかの意見だけを採用する。	相手と自分の意見が合わなければ、調整しようとは何かの提案をするが、独りよがりな提案である。	みんなが納得するように話し合えるような促進ができる。
応答性	働きかけに無反応。	反応はするが、文脈には合っていない。	ある程度適切な反応がわかっているが（文脈に合っているが）、適切でない。	相手からの働きかけに反応するが、やりとりが続かない。積極的ではない反応だが、論点はずれていない。反応するが、結果的にできていない。意味に応じた反応をしようとしているが、十分でない。	相手の働きかけに適切な反応・言葉の対応ができる。気持ちよい。相手の話の意味内容に応じて反応できている。
頼む	頼むという行動が見られない。	何か頼んでいるようだが、はっきりと言葉にして頼めていない。	頼みたい事は言うが、その理由を言ったり、かなえられたらどうなるかを言えていない。	理由をつけて頼み事を言える。ただし、かなえられたらどうなるかを言えていない。	理由をつけて頼み事を言い、またそれがかなえられたらどのようになるかを言える。

Table 2 Rubricの「具体例・評定のポイント」の例（応答性スキル）

	Not yet 全く見られない 1	Novice 初心者程度 2	Apprentice まああ 3	Proficient だいたいできている 4	Distinguished 完璧 5
具体例・評定のポイント	他者に無関心。働きかけに無反応、無視、気づかない。声をかけても反応がない。	あざけり。うつむいて目を合わせない。ボソボソと相手にわかりにくい反応。求められていることとは違う反応。文脈に関係なく「いやだ」。	トンチンカン。愛想がない。ぶっきらぼう。「あっ、そう」、「別に」、「わからない」、「しらない」。愛想なくぶっきらぼうに「いやだ」（ただし文脈には合っている）。	相手との話を継続させるような反応ではない。「いやだ」ときちんと言意は伝えるが、会話は発展しない。	話の流れに沿った受け答えが来ている。相手の意見を引き出すような質問。相手の話しかけに言葉やNVでの反応を見せ、相手に対して話を継続、発展させるようなやりとりをすることができる。

Table 3 行動チェックリスト

スキル	項 目
話す	グループの話し合いの中で、自分の意見をはっきりと話した
	ちゃんと理由をつけながら自分の意見を言った
聞く	授業をよく聞いた
	グループでの話し合いで、人の話をよく聞いた
協力	グループで協力しようとした
	人のよいところに気づき「すごいね」などと言った
調整	人の意見を尊重した行動をとった
応答性	他者からの働きかけに、言葉で適切な受け答えをした
	他者からの働きかけにうなづくなど、NVで適切な反応をした
	「ありがとう」と言った
	人にほめられたら、素直に喜んだ
頼む	気持ちよく頼みごとをしていた

するというTaskであった。話すスキル、聞くスキル、協力スキル、調整スキル、応答性スキルが観察できるように、情報カードを1枚受け取るたびに推理のための話し合いをさせたり、解答を全員で確認しながら地図に記入させたり、答え合わせのときの発表者を相談で決めさせたりした。また、地図の余白に自由に書かせることによって、さらにこれらのスキルを観察する機会を増やした。また、頼むスキルを観察するために、グループの全員が順に異なる頼み事を書かれたカードに従って頼み事を行うとそのたびに新しい情報カードを受け取れるルールにし、頼む場面を意図的に課した。記録は、Task1で対象児全員を画面に捉えきれていなかったことと音声不明瞭だったことから、ビデオカメラ各2台と集音マイクを使用することでより詳細に録画・録音された。

3. Rubricに基づく評定

- ①評定者：活動立案と実施者であるスタッフのうち中心的なスタッフ9名（うちRubric作成担当者5名）。グループ1を4名、グループ2を5名のスタッフが、Task1、Task2を通して同じ対象児を評定した。以下、この評定のそれぞれを評定1、評定2と表記する。
- ②評定日：評定1は12月5日、評定2は12月12日に行われた。
- ③評定の手続き：評定者全員がRubric作成担当者から評定方法の説明（Rubricの説明、録画画像の観察回数制限はないこと、巻き戻しての観察可など）を受けた後、Taskが録画されたDVD（加えて、評定1では音声ファイル）、Rubric一覧表、評定記入表が配付され、Rubric一覧表に基づき、最頻行動評定（Task中に最も多く見られた行動に該当する水準）と最高行動評定（Task中に見られた最高の行動に該当する水

準）が求められた。なお、評定1を実施した結果、話すスキルと聞くスキルのRubricに若干の不適切な点が評定者によって指摘された。また、調整スキルと応答性スキルでは、Rubricの不明確さにより評定が不可能と判断した評定者がいた。そこで、これらのスキルのRubricを改訂してから評定2が実施された。このように評定1ではRubricの不備により良好なデータを得られなかったため、評定1と評定2を厳密に比較することができない。ただし、Rubricの改訂のため、以後の分析で評定1のデータも比較対象として活用した。

4. コミュニケーション行動チェックリスト

わくわくコミュニケーションクラブ秋クラスでは、全6回を通して、活動が終わるたびに、各グループの担当スタッフ2名が、本活動で子どもに獲得させたい16項目のコミュニケーション行動について、活動中に自分が担当しているグループの各児童の活動の中に1回でも見られたかどうかをバイナリーチェックした。本研究では、この6回分の行動チェック数をRubricによる評定との関連を検討するための資料として用いた。ただし、本研究で用いる行動チェックリストの項目は、Rubricと対応している12項目とした（Table3）。

結果と考察

1. 評定者間の評定の安定性

本研究で作成したRubricに基づく評定においては児童1人に対して4人、もしくは5人のスタッフが評定者となった。この評定が一致する程度を確認するために、児童ごとに標準偏差を算出した（Table4）。

評定1では、話すスキルと聞くスキルは、最頻行動評定、最高行動評定とも標準偏差が大きかった。つまり、

Table4 対象評定児ごとの評定値の標準偏差

児童	＜評定1＞最頻行動評定					＜評定1＞最高行動評定					＜評定2＞最頻行動評定					＜評定2＞最高行動評定							
	話す	聞く	協力	調整	応答性	話す	聞く	協力	調整	応答性	話す	聞く	協力	調整	応答性	頼む	話す	聞く	協力	調整	応答性	頼む	
グループ1	1	.50	.43	.50	.43	.43	.87	.43	.83	.00	.50	.43	.00	.00	.43	.50	.43	.43	.71	.50	.43	.50	.43
	2	.00	.83	.00	.43	.83	.43	.50	.43	.71	.43	.00	.00	.00	.43	.50	.50	.71	.71	.00	.43	.50	.50
	3	.50	.43	.43	.43	.50	1.12	.43	.43	1.12	.50	1.00	.71	.00	.00	.50	.00	1.22	.50	.71	.43	.50	.71
	4	1.12	1.00	.00	.00	.43	.87	.50	.50	.43	.50	1.00	.43	.43	.71	.43	.43	.50	.50	.00	.43	.50	.83
	5	-	-	-	-	-	-	-	-	-	-	.00	.83	.00	.83	.00	.50	.43	.50	.50	.83	.43	.00
	6	-	-	-	-	-	-	-	-	-	-	.43	1.00	.00	.00	.00	.43	.50	.43	.50	.43	.43	.43
グループ2	7	.63	.98	.40	#	#	#	.75	.00	#	#	.63	.00	.00	1.26	.40	.80	.49	.49	.49	.98	.40	.00
	8	.75	.75	.49	#	#	.49	.75	.00	#	#	1.10	.80	.00	.98	.49	.40	.40	.49	.63	1.02	.49	.80
	9	.80	.63	.80	.80	.75	.75	1.02	.75	.00	.75	-	-	-	-	-	-	-	-	-	-	-	-
	10	.49	.63	.49	.49	#	.49	.75	.40	.80	#	.80	.49	.49	.80	.00	.98	.49	.00	.40	.75	.40	.49
	11	1.17	.89	.40	.80	.40	.75	.75	.49	.80	.40	.00	.40	.40	.75	.40	.63	.40	.40	.00	.49	.00	.49

■ : SD=.50 □ : SD=.00 - : 欠席のためデータの欠如 # : 評定者の一部が評定不可能と判断

話すスキルと聞くスキルの評定が評定者間で不一致の度合いが高かったということになる。また、調整スキルと応答性スキルを評定不可能と判断した評定者がいたが、これは、調整スキルを観察できる場面がTask中に少なかったことやRubricに不備があったためと考えられる。協力スキルは、標準偏差が小さいものが多く、評定者の評定もいくつか完全一致していた。また、応答性スキルに評定不可能データもあるものの、標準偏差が小さいものが多かった。つまり、これらのスキルについてはRubricに基づく評定に一定の信頼性があったと言うことができよう。

ところで、評定1で明らかとなったRubricや評定に関する問題点を改善するため、①話すスキルと聞くスキ

ルの評定水準の明確化、②調整スキルと応答性スキルの具体例の充実、③録画画像と音声の記録方法の変更を行った上で評定2を実施したところ、話すスキルや聞くスキルは最頻行動評定、最高行動評定ともに標準偏差が小さくなり、特に最頻行動評定では評定者の評定が完全一致する児童が多くなった。また、協力スキル、応答性スキル、頼むスキルについても、標準偏差は小さく、評定者間の評定も一致する傾向であったことから、評定2では、Rubricによる評定がより信頼できるものに改善されたと言える。

本研究の評定者は、評定前に特別に訓練を受けていなかったため、評定1が結果的に訓練の意味合いを担うこととなり、評定2でより正確な評定ができるようになった。

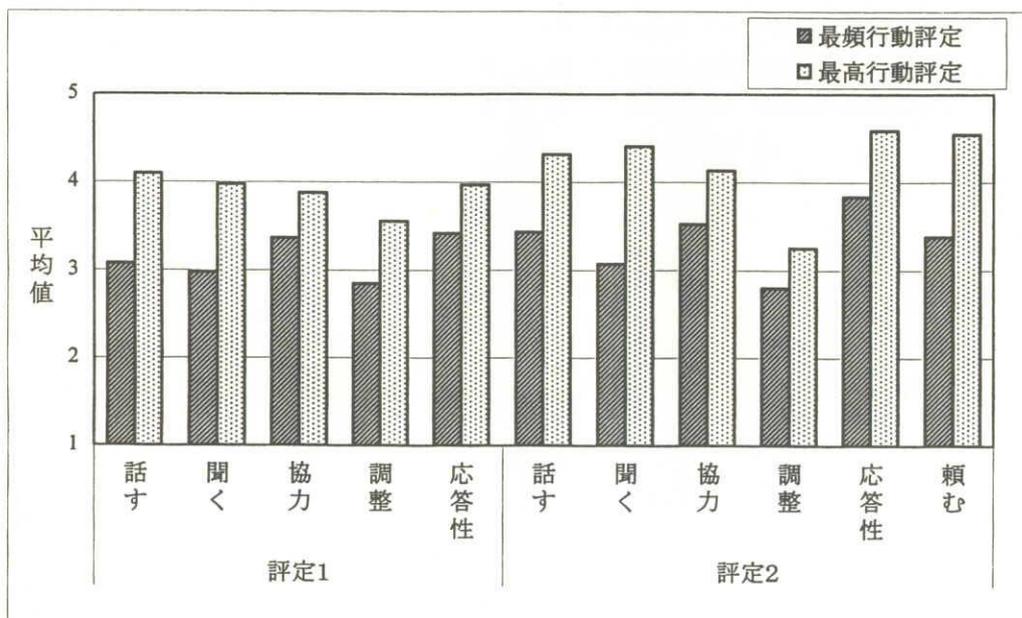


Figure 1 各スキルの評定平均値 (評定1、評定2)

たということも、評定の一致度を高めた重要な要因であると考えられる。PAを行う際には、評定者の事前の訓練が、信頼性の高い評価を行う上で重要なポイントになってくると考えられるであろう。

2. 最頻行動評定と最高行動評定の差

最頻行動評定および最高行動評定の平均値を各スキルごとに算出した (Figure1)。

最頻行動評定、最高行動評定のいずれにおいても調整スキルの平均値が他のスキルに比べて低い。これは Task1、Task2 ともに、意見が対立するような対人葛藤場面がほとんどなく、調整する機会が少なかつたためであると考えられる。また、調整以外の他のスキルの平均値は3以上であった。3というのは、「まあまあできている」ことを意味するため、対象児のコミュニケーションスキルは、「まあまあできている」水準より上であると言つてよい。

なお、各スキルの最頻行動評定と最高行動評定の差を確認するために、評定1、評定2の各スキルにおける最頻行動評定と最高行動評定の差の検定を行った結果、すべてのスキルにおいて有意な差が見られた。

Table 5 評定2における最頻行動評定と最高行動評定の相関

	最頻行動評定					
	話す	聞く	協力	調整	応答性	頼む
最高行動評定	.67 **	.28	.59 **	.79 **	.31*	.11
	*5%水準	**1%水準	.4 以上			

3. 最頻行動評定と最高行動評定の相関

評定2における各スキルの最頻行動評定と最高行動評定の相関係数を算出し、両者の関連について検討を加えた (Table5)。なお、評定1は評定不可能と判断した評定者がおり、データが不完全なため、この分析からは除

外した。

話すスキルと協力スキル、調整スキルで最頻行動評定と最高行動評定の相関が高かつた。一方、聞くスキルや応答性スキル、頼むスキルでは弱い相関しか見られなかつた。これらのスキルの最高行動評定の平均値は高く、その標準偏差も小さいことが一因であるが、そういった相関の弱さが、最頻行動と最高行動との質的な違いを意味するのか、それとも、評定の対象となつた活動の質によるものなのか、さらには、評定の対象となつている子どものスキルの高さによるものなのか、そのいずれかを客観的に同定することは現時点では不可能である。活動Taskの種類、子どものスキルのバリエーションなどを拡大した検討が必要である。

4. 最高行動評定と最頻行動評定の差についてのスキル間の比較

最高行動評定と最頻行動評定との差 (持っているスキルが活動中に常に発揮されている程度) がスキルによって異なることが予想された。そこで、評定2におけるスキルの種類を独立変数、最高行動評定値と最頻行動評定値の差を従属変数とした1要因分散分析を行った。なお、評定1はデータが不完全なため、分析からは除外した。

その結果、有意な主効果が見られた ($F(5, 258) = 4.91, p < .001$)。そこで、スキル間の差を確認するため、多重比較を行ったところ、聞くスキルと協力スキル、聞くスキルと調整スキル、聞くスキルと応答性スキル、協力スキルと頼むスキル、調整スキルと頼むスキルの間に有意な差が見られた (Figure2)。つまり、聞くスキルや頼むスキルは、自分の持っている最高水準のスキルが頻繁に発揮されにくい、協力スキルや調整スキル、応答性スキルは、頻繁に発揮されやすいスキルと考えることができよう。

これらのことから、日常では持っているスキルの最高水準を頻繁に発揮しているスキルとそうでないスキルがあることが示唆された。ただし、こういった考察も、他

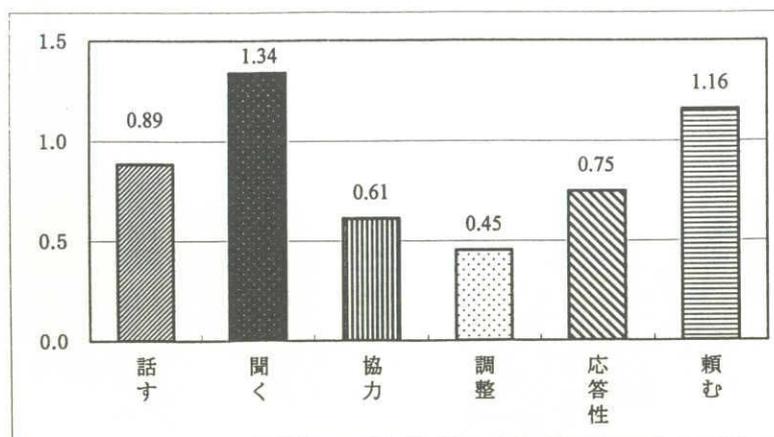


Figure 2 最高行動評定と最頻行動評定の差についてのスキル間の比較 (評定2)

Table 6 評定2と行動チェックリスト得点の相関

	Rubric											
	話す		聞く		協力		調整		応答性		頼む	
	最頻	最高	最頻	最高	最頻	最高	最頻	最高	最頻	最高	最頻	最高
行動チェックリスト	.47	.61	.30	.75*	.19	.03	-.60	-.34	.30	.59	.40	.50

*5%水準  |.4|以上

の性格を持った活動Taskを用いた検討を待たなければならぬことは言うまでもない。

5. Rubricによる評定と行動チェックリスト得点の関連

本研究で取り上げたスキルのRubricに基づく評定と、行動チェックリストの得点（1回あたりの活動でスキルを用いた行動が見られたかどうか）との関連性を検討するために、評定2と、行動チェックリスト得点の相関係数を算出した（Table6）。分析の際、Rubricに基づく評定、行動チェックリストの得点ともに児童1人あたりの評定者の平均値を用いた。また、行動チェックリスト得点は、全6回分の合計を算出し、1回あたりの平均値を用いた。なお、Rubricに基づく評定における評定1はデータが不完全なため、分析から除外した。

評定2における話すスキルと頼むスキルの最頻行動評定、最高行動評定とも、行動チェックリスト得点との相関は、調整スキルを除いて比較的高かった。つまり、Rubricによって評定された話すスキル、頼むスキルと行動チェックリストでチェックされた話す行動と頼む行動の得点に対応していることになる。

ただし、調整スキルの相関は、最頻行動評定、最高行動評定ともに負の相関であった。これは、行動チェックリストの調整スキル項目（Table3）とRubricの調整スキルが異なるもの、もしくは逆のものを指していることによる。つまり、行動チェックリストの調整（相手の意見を尊重する）は、必ずしもRubricの調整（意見があわないうきに折り合いをつけて、納得できる提案をする）と同じではなく、納得できる提案をするという比較的高度なスキルがRubricにおいて評価されていたからだと考えることができるのである。さらに、Task2では調整スキルを発揮する場面が少なかったため、評定者によって評定が一貫しなかったためとも考えられるが、評定の際のスキルの定義という操作的な定義が異なっていることが相関を検討することには大きな意味を持たないことを考えれば、Rubricによる調整スキルの妥当性について議論することは尚早である。

また、協力スキルに関しては、最頻行動評定、最高行動評定ともに相対的に相関が低かった。つまり、Rubricによって評定された協力スキルと行動チェックリストでチェックされた協力スキルが対応していない。しかし、

行動チェックリスト得点は全6回活動の平均値であるため、各回の活動において協力行動の表出が一貫していなかった可能性も考えられる。

さらに、全般的には、最高行動評定において行動チェックリスト得点との相関が高い傾向が読み取れる。Rubricの最高行動評定とは、Taskの中で見られた最高水準のスキルの評定値であり、行動チェックリスト得点は活動の中で1回でも見られた行動がチェックされたものの総和である。こういった各資料の持つ特徴を考えれば、こういった結果が析出されることは、最高行動評定と最頻行動評定の2指標を独立にとっていくことの意味を物語っていると考えることができる。

総合考察

1. PAに基づいたRubricとTaskの開発

われわれの2年間にわたる子どものコミュニケーション能力を向上させるための実践研究においては、その効果測定の方法が最大の課題であったが、本研究では、試行的にPAという手法を用いて行動観察によるスキル評定の開発を試みた。

1) 評定者の視点の変化と問題点

評定1では同一児の評定値の標準偏差が大きかったが、この評定1で評定者が評定訓練をできたことと、評定者のPAに対する積極的態度を得てRubricの不備を改善できたことにより、評定2において評定者の評定の信頼性を高めることができたと言えよう。また、評定はわくわくコミュニケーションクラブの活動と同時に進んだが、評定者がスタッフも兼ねていたために、評定1を実施した後のTask2では、評定者がRubricを念頭に置いて子どもを観察し促進的な援助ができたという副次的効果も得られた。Rubricの存在が、スタッフとしての教育視点を意識させ、子どもへの援助行動を導いたという、援助の好循環を生み出したことは貴重な収穫であった。しかしながら、活動スタッフが評定を行う問題も同時に指摘できよう。スタッフは各回の活動で評定対象児に関わっているため、すでに児童のスキルの表出に関する認知的枠組みを持っており、それが評定に影響している可能性も十分に考えられるためである。今後は、評定の信頼性を確保するためにも、活動に直接携わっていない研究協

力者による、ある程度の評定訓練後の評定を得て、その評定とスタッフの評定との比較をする必要がある。

2) Rubric の開発

本研究で注目した6つのスキルについて、評定する際の抛り所となる Rubric の開発に取り組み、その Rubric の信頼性を、評定者間の評定値の一致度によって検討したところ、評定1では同一児の評定値の標準偏差が大きかったため、Rubric を改訂して評定2に臨み、一定の成果を得ることができた。評定前の訓練も含め、評定者に Rubric と評定方法をよく理解させることがこの種の評定・評価には不可欠である。

また、行動チェックリスト得点と Rubric 得点の相関を検討したところ、データ数は不十分ながら、話すスキル、聞くスキル、応答性スキル、頼むスキルにおいて相関関係が見られた。行動チェックリストの項目の表している内容が、断片的であったり、Rubric の内容と合致していなかったりしたため、今回は、行動チェックリストとの相関によって妥当性を検討することはできないが、今後は行動チェックリストを改訂することによって、さらには PA とは異なった他者評定や自己評定などを平行して実施することによって、Rubric の妥当性を検証していく必要がある。また、これらの問題を解決しながら、より多くのスキルについて Rubric を開発していくことも必要である。しかし一方で、妥当性と信頼性を確保するために要する時間と労力はかなり大きいのも事実である（鈴木，2004、松下，2005）。

3) Task の開発

各スキルを評定するための Task の開発も Rubric と同様に重要である。Task2 では、短時間で被験児全員を観察することが難しいと判断された頼むスキルを観察できるように意図的に場面設定をしたが、子どもは Task2（秋クラス第6回活動）よりも以前に「頼み方（秋クラス第3回活動）」を学習したことと、ゲーム中の課題だったことなどが影響し、戸惑うことなく頼み事ができていた。中には得意げに頼み事ができた児童もいた。制作活動、作業活動などをゲーム的に組み入れた Task を設定することによって、日常に近い状態でターゲットスキルを観察できることが確認されたため、スキルトレーニングのアセスメントとしてこのような Task が有効であるという感触を得ることができた。しかし、調整スキルについては、今回の Task では観察できる場面が少なかったことも事実である。今後は評定するスキルを十分に観察できるような Task を開発していく必要がある。ただし、Task 開発にもまた多大な時間と労力が要求されるため（鈴木，2004、松下，2005）、スキル測定に特化した一定の定式化された Task を準備する必要があるのかもしれない。

2. スキルの般化と評価の活用

話すスキル、協力スキル、調整スキルの最高行動評定と最頻行動評定の相関が高いことから、高い水準の行動スキルを持っている子どもは日頃の行動スキルも高いことが推測される。よって各スキルの最高行動評定を高めるような働きかけをすることによって行動が日常に般化される可能性が示唆されたと考えられることも無理ではないだろう。また、最頻行動評定と最高行動評定の差の、スキルによる違いを検証したところ、特に聞くスキルと頼むスキルについて、持っている最高水準のスキルを必ずしも常に表出しているわけではないという結果であった。しかし、例えば頼むスキルの最高行動評定である「頼むときの3つのコツ（廣岡ら，2005a）」を日常で毎回表出することは、必ずしも自然なことではない。なぜなら、子どもの日常における社会的場面においてそのようなスキルを子どもが発揮することが、必ずしも社会的には適切ではないことが多くあるからである。このように、それぞれの社会的場面で望まれるスキルの水準が異なるとすれば、最高行動、最頻行動を分けて評定することには十分な意義があると考えられる。

3. ソーシャルスキルの評価法としての PA

ソーシャルスキルの測定・評価としては、自己評定法、他者評定法（役割演技法、面接法、自然観察法など）などが一般的に知られている。しかし、小学生のソーシャルスキルの客観的測定という目的からすると、小学生という発達段階にある子どもに自己認識能力が十分に発達しているとは言えず、自己評定法によって客観的に測定することは難しいと言わざるを得ない。また、仮に発達の問題がないとしても、測定の対象が特定の場面や課題にのみ限定されてしまい、一般化が困難であるといった問題も指摘されている。一方、自然観察法は、観察者による行動評定に依存しているので、発達段階の低い小学生のスキルの測定には最適な方法だと考えられるが、行動が自然に生起する可能性を考えると、測定に対する時間的コストはかなり高くなる。こういった認識から、日常的に起こりうる自然な文脈を維持しながら、その中に現実的な相互作用場面を Task という形で設定し、いわば実験的な場面における子どもの行動を観察することができる PA という評価法を、われわれの活動の効果測定として選択したわけである。

PA は我が国においては新しい教育評価法である（Caroline, 2001; 鈴木, 2004）。先述したように PA を実現するための課題は大きいですが、具体的な行動を他者が観察して Rubric に基づき評価するといった手続きを踏むことは、そのコストをかけるに値する利点があり、スキルトレーニング等の効果測定として妥当なものであろうという感を強くすることができた。さらに、コミュニケー

ション能力の測定といった視点からPAがより広範囲な社会的能力の測定に適用可能であることが示されれば、今後の社会性教育実践場面において活用される可能性は高い。

引用文献

- Caroline V. G. 鈴木秀幸訳 2001 新しい評価を求めて—テスト教育の終焉 論創社
- 廣岡秀一・元吉忠寛・小川一美・斎藤和志 2001 クリティカルシンキングに対する志向性の測定に関する探索的研究(2) 三重大学教育実践総合センター紀要, 21, 93-102.
- 廣岡秀一・中西良文・廣岡雅子・後藤淳子・横矢規・矢神祥代・福田真知 2005a 小学生のコミュニケーション力を高める教育実践 三重大学教育学部研究紀要(教育科学), 56, 317-332
- 廣岡秀一・中西良文・廣岡雅子・後藤淳子・横矢規・矢神祥代・福田真知 2005b 小学生のコミュニケーション力を高める教育実践(2) 三重大学教育実践総合センター紀要, 25, 37-45
- 廣岡秀一・中西良文・廣岡雅子・横矢祥代・福田真知・秋山美和・伊藤由恵 2006 小学生のコミュニケーション力を高める教育実践(3) 三重大学教育学部研究紀要(教育科学), 57, 101-110
- 廣岡秀一・小川一美・元吉忠寛 2000 クリティカルシンキングに対する志向性の測定に関する探索的研究 三重大学教育学部研究紀要(教育科学), 51, 161-173.
- 松下佳代 2005 学力評価法の新たな可能性 日本教育心理学会第47回大会発表論文集, 47, 18-19
- 鈴木秀幸 2004 「新しい評価と思考力、判断力、意欲」指導と評価, 50, 4-8
- 吉田俊和 1997 対人関係能力の低下が「社会」にもたらす影響(討論論文2) 誌上シンポジウム「対人関係の低下と現代社会」 名古屋大学教育学部紀要(心理学), 44, 29-32
- 吉田俊和・廣岡秀一・斎藤和志(編) 2002 教室で学ぶ「社会の中の人間行動」—心理学を活用した新しい授業例— 明治図書 p107