

第27回 第二言語習得研究会全国大会

2016年12月17日（土）九州大学

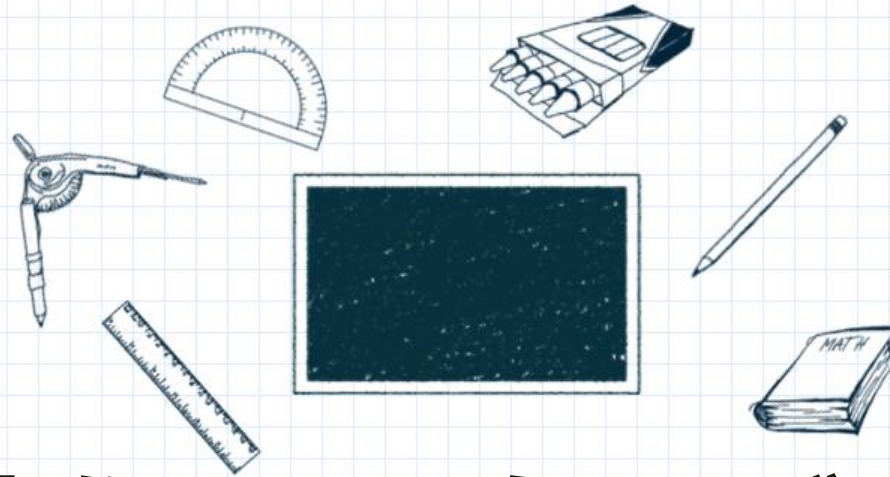
パネルディスカッション：「第二言語習得における口頭運用能力とは－理論・実践・評価－」

口頭運用能力の評価法

－「妥当性・真正性」再考－

広島大学 渡部 倫子

tomokow@hiroshima-u.ac.jp



「唯一の正解」がない

口頭運用能力をどう評価すれば？

1. 真正性の高い評価法**も**用いる。
2. 実施前後に妥当性を検証する。
3. 結果を共有する。

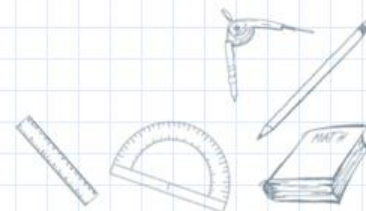


言語評価研究の目的

(Bachman, 1990: 13)



1. テストによって言語運用能力の理論を構築・検証
2. テストの実施・結果の解釈が妥当かを確認
3. テスト結果と実世界での言語使用との関係を調査

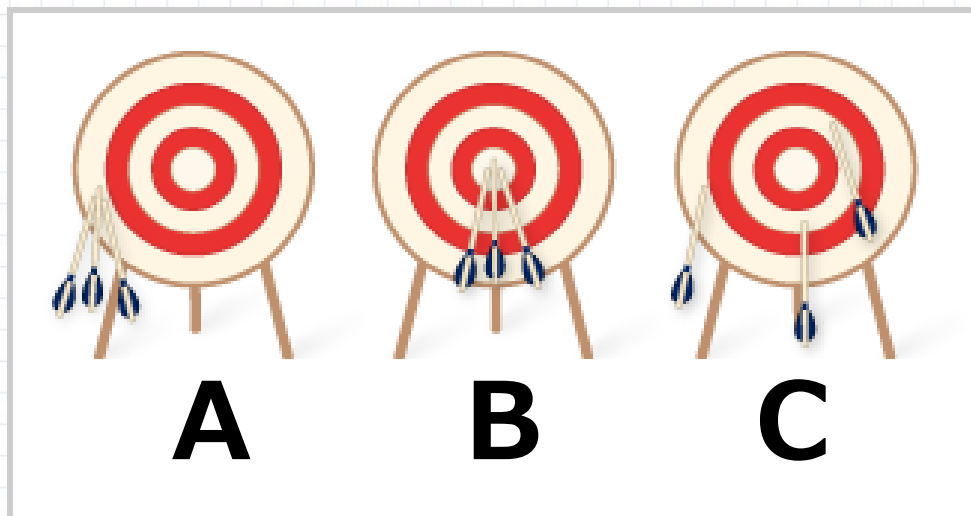




妥当性と真正性



妥当性が高い評価法とは？（矢＝採点者）



<http://blog.wonderlic.com/reliability-and-validity-in-testing-what-do-they-mean>

A：信頼性は高いが妥当性が低い。

B：信頼性・妥当性が高い。

C：信頼性が低く、従って妥当性も低い。

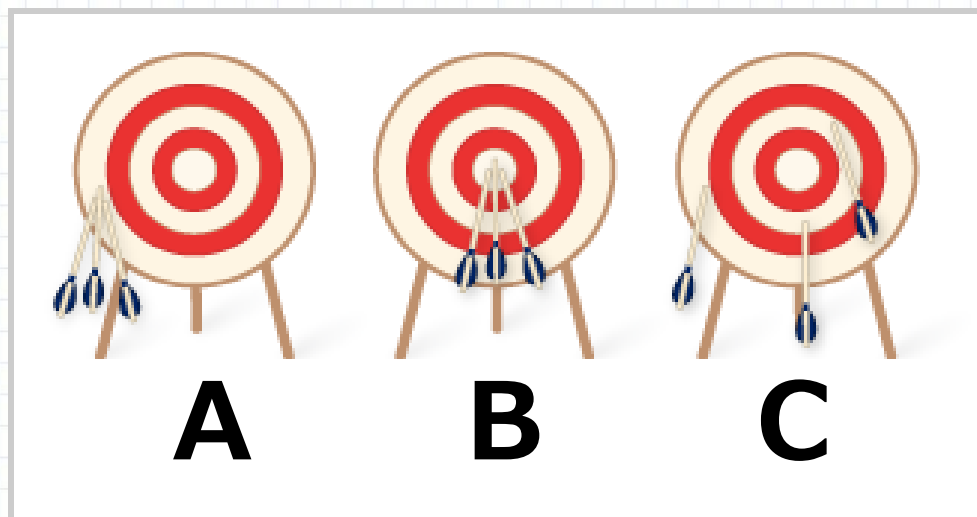




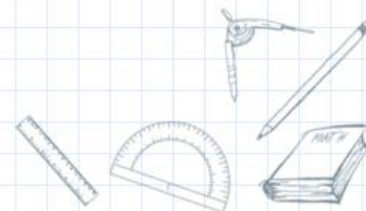
妥当性と真正性



真正性が高い評価法（＝弓道の段位試験）
とは？



<http://blog.wonderlic.com/reliability-and-validity-in-testing-what-do-they-mean>





妥当性と真正性



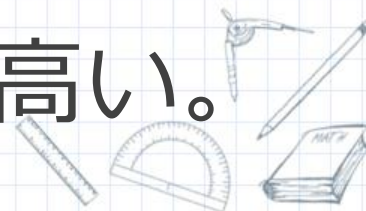
弓道の段位試験のペーパーテストより、

1. 競技に使用できる弓の長さは、何
cmが基準となっているか。

() cm

2. 「残心」とは何か説明しなさい。

実技試験のほうが真正性が高い。





ペーパーテストよりも真正性が高い 口頭運用能力の評価法



1. モノログ（1人）

- ① 声に出して読む（音読）
- ② スピーチ
- ③ 絵・写真・実物描写
- ④ あらすじ・要旨の口頭再生



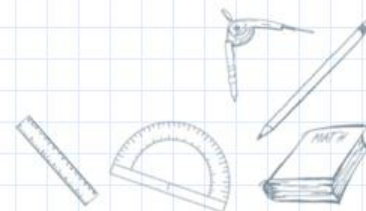
2. ダイアログ（2人）

- ① インタビュー（面接）
- ② ロールプレイ



3. マルチログ（3人以上）

- ① ディスカッション
- ② ディベート





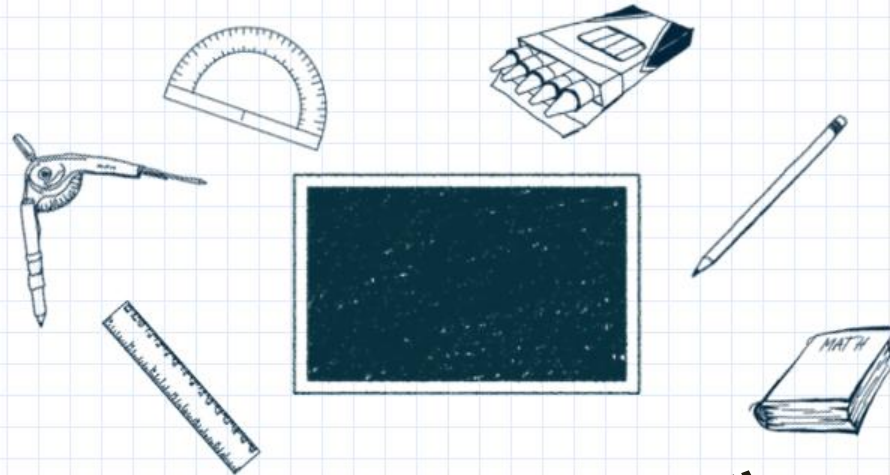
評価法の有用性 (Bachman, 1990)



1. 妥当性：得点＝構成概念と解釈できる？
2. 真正性：実生活で行う課題？
3. 信頼性：結果が一貫している？
4. 相互性：学習者の特性との関わりは？
5. 影響：社会、機関、個人への影響は？
6. 実用性：実施が可能？

妥当性・真正性以外にも考慮すべき有用性の特質がある。





妥当性と真正性のバランスは どうとれば？

1. 評価の目的によって異なる。
2. 論証モデルを用いて、評価の実施前後に検証する。



論証モデルによる真正性の検討

(Bachman and Palmer, 2010; Kane, 2006)



解釈的論証

根拠

Aの得点が
高い

主張

Aの能力は
高い

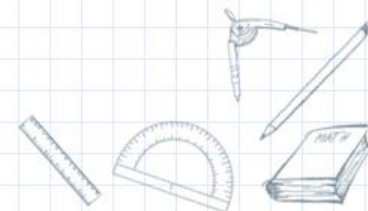
理由

面接は
真正性が高い

裏付け

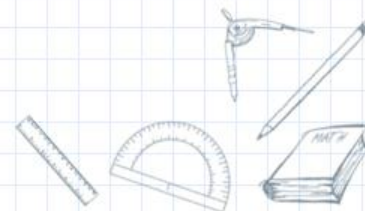
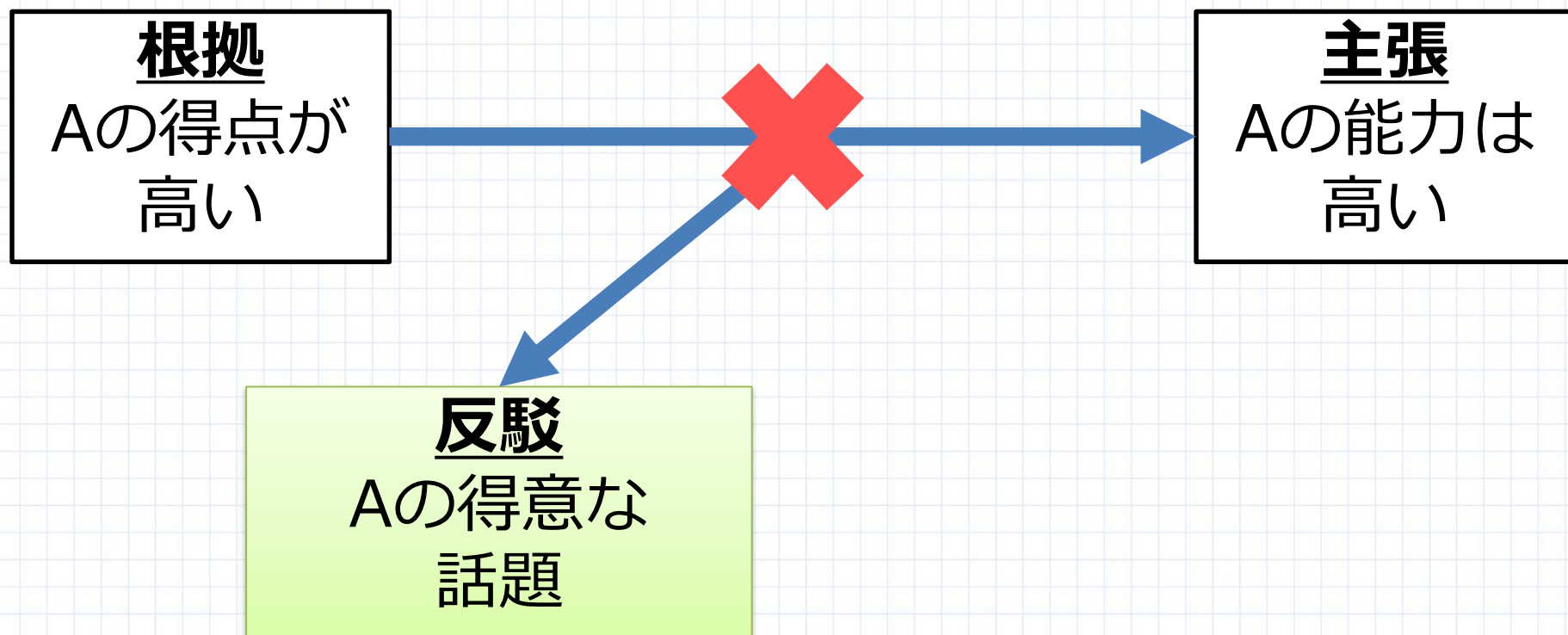
Aは面接を
受ける予定

妥当性の論証





論証モデルによる相互性の検討





論証モデルによる相互性の検討



解釈的論証

根拠

Aの得点が高い

主張

Aの能力は高い

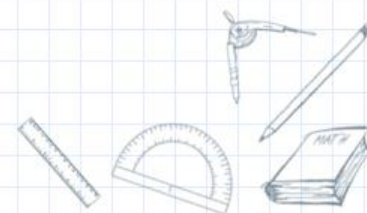
反駁

Aの得意な
話題

却下の裏付け

複数の
話題を使用

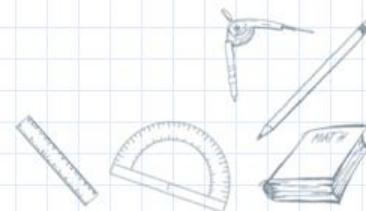
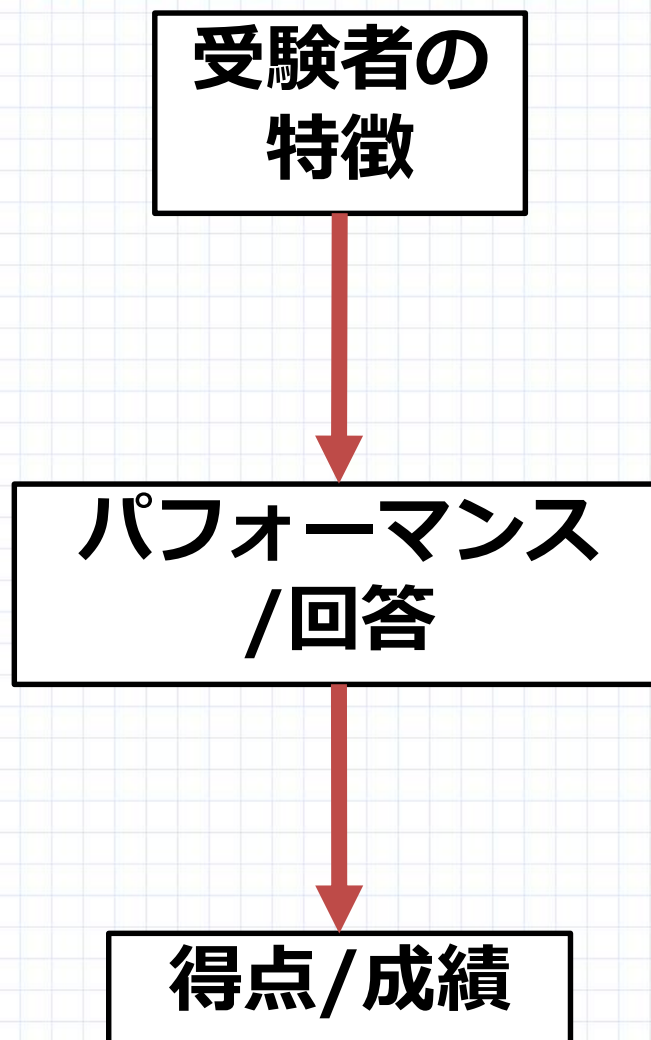
妥当性の論証





社会的・認知的枠組みによる妥当性の検証

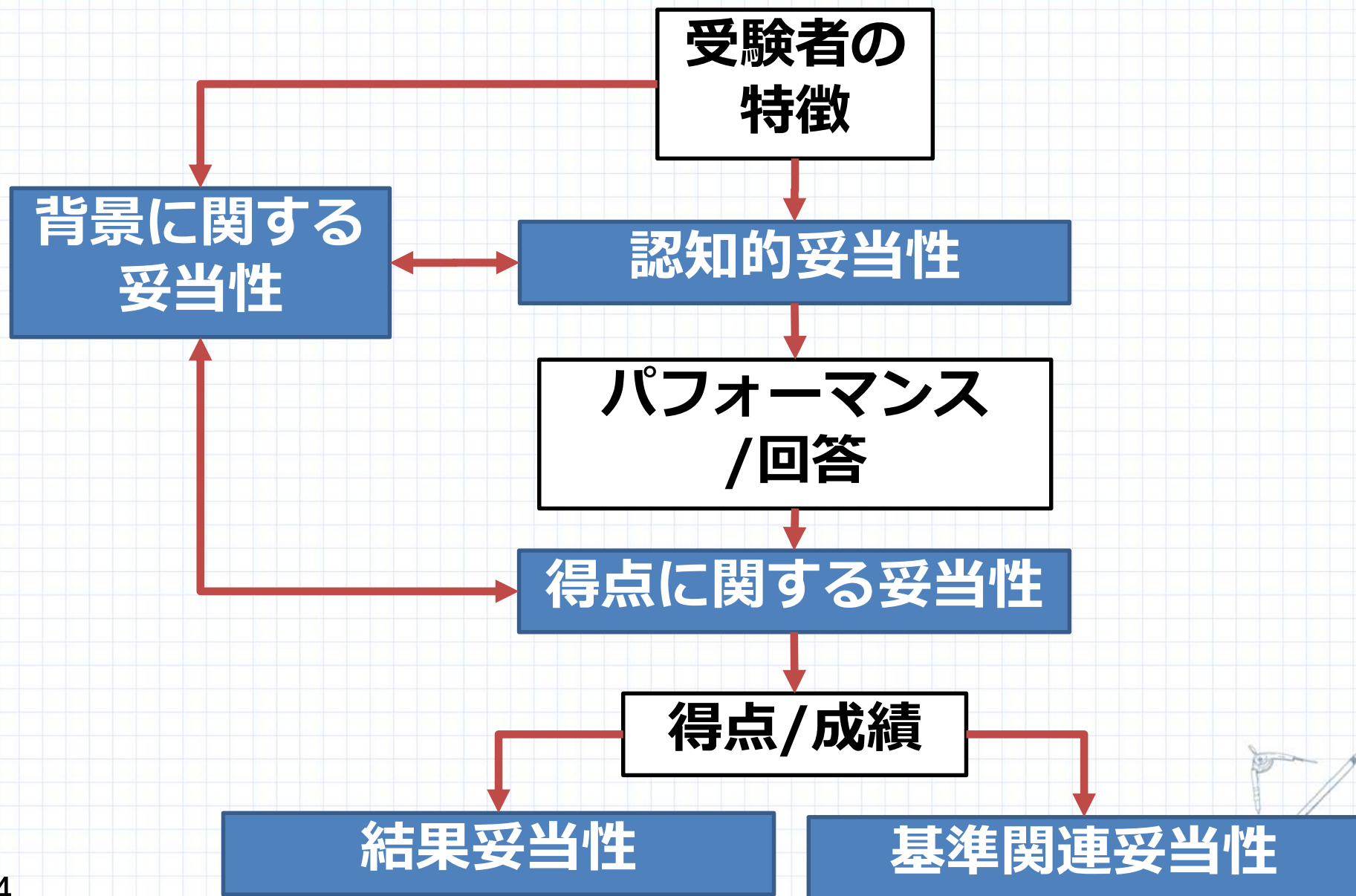
(Weir, 2005)

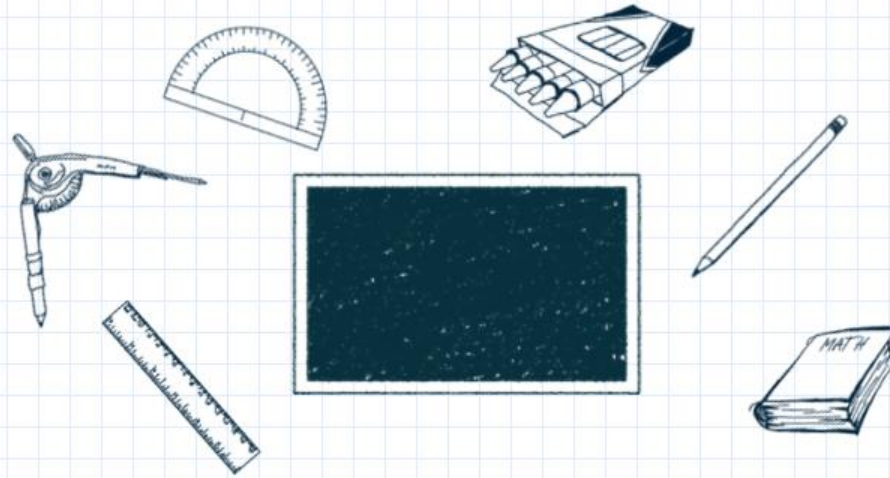




社会的・認知的枠組みによる妥当性の検証

(Weir, 2005)



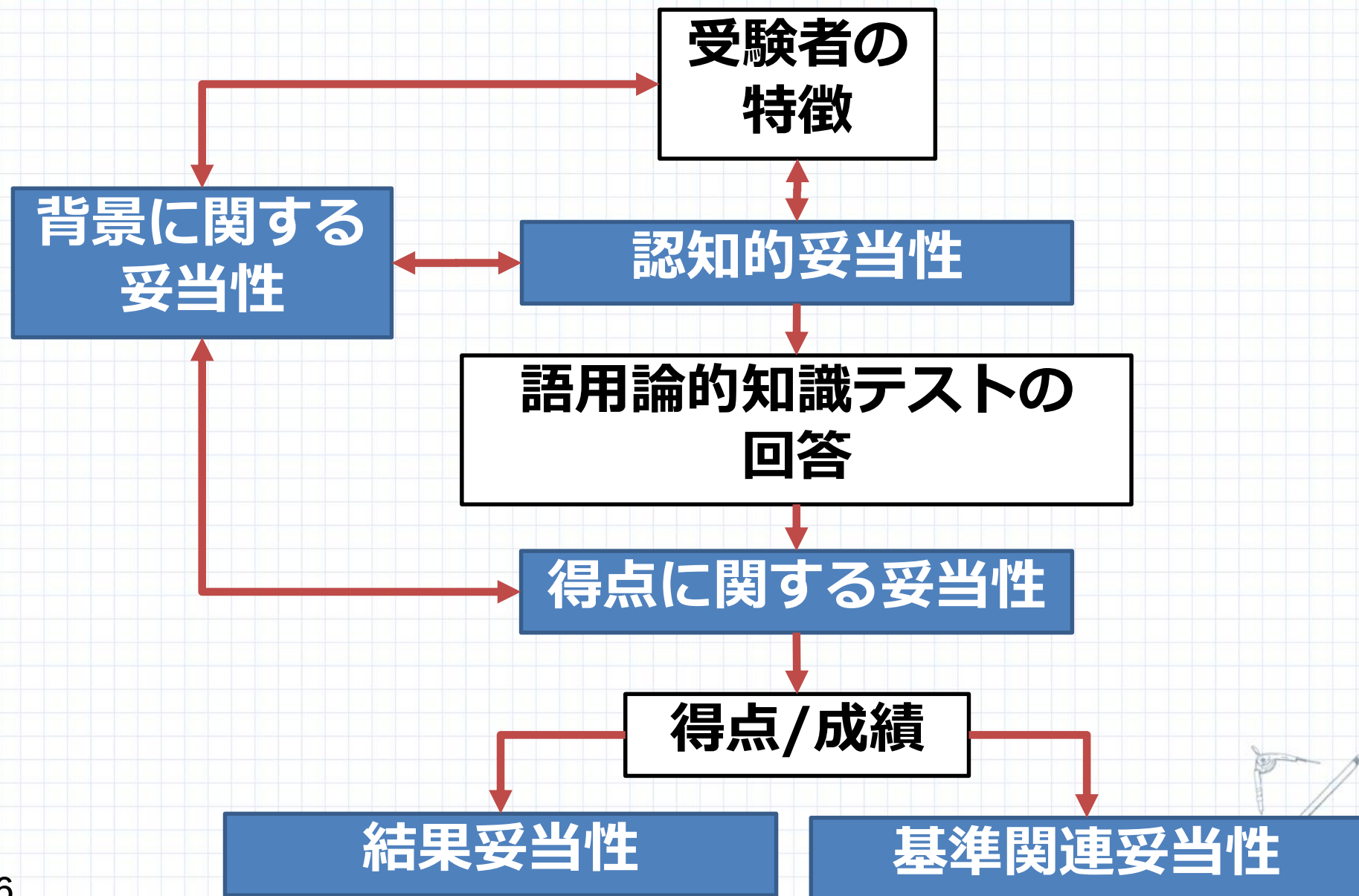


論証モデルと社会的・認知的枠組み どちらを使えば？

1. 社会的・認知的枠組みを用いる。
2. 論証モデルを併用する。
3. 必要な裏付けを提示する。



語用論的知識テスト(糸満, 2011) の検証





語用論的知識テスト (糸満, 2011: 29)



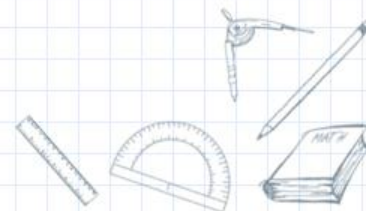
- **多肢選択問題**：48問（発話行為，決まり文句，スピーチスタイル）

例）客が家電店で店員と話しています。客は店員に何か頼んでいます。客は店員に何と申うでしょうか（発話行為）。

「すみませんが，ちょっとそれ_____。」

- A. 見せてくれませんか
- B. 見せたいんでしょうか
- C. 見せてあげましょうか
- D. 見せてもらいたいでしょいか

- **正誤を得点化**
- 日本語学習者119名





語用論的能力の評価法

(石原・コーエン, 2015; Rover, 2011)

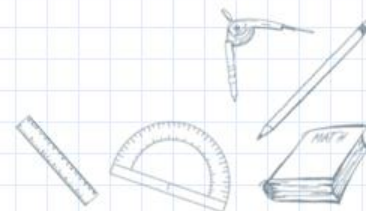


1. 理解（筆記・リスニング）

- ① 評価スケール問題：依頼の適切さの高低
- ② **多肢選択問題：依頼表現から選択** (糸満, 2011)
- ③ ランク付け：依頼の気軽さランキング
- ④ 自由記述：①～③の理由

2. 産出

- ① インタビュー
- ② ロールプレイ
- ③ DCT（口頭・記述・ビジュアル、目的選択）
- ④ 短文完成問題
- ⑤ 学習者によるデータ収集・考察





語用論的能力の得点化・フィードバック

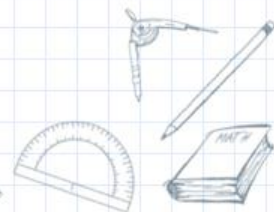
(石原・コーエン, 2015; Rover, 2011; Youn, 2015)

1. 理解（筆記・リスニング）

- ① **正誤を得点化**(糸満, 2011)
- ② 数値的な尺度：適切さの4段階尺度
- ③ コメント

2. 産出

- ① 数値的な尺度
- ② ルーブリック：数値的な尺度＋記述
- ③ コメント
- ④ 会話分析・談話分析結果の提示





ループリックの例

(Youn, 2015: 223-224)

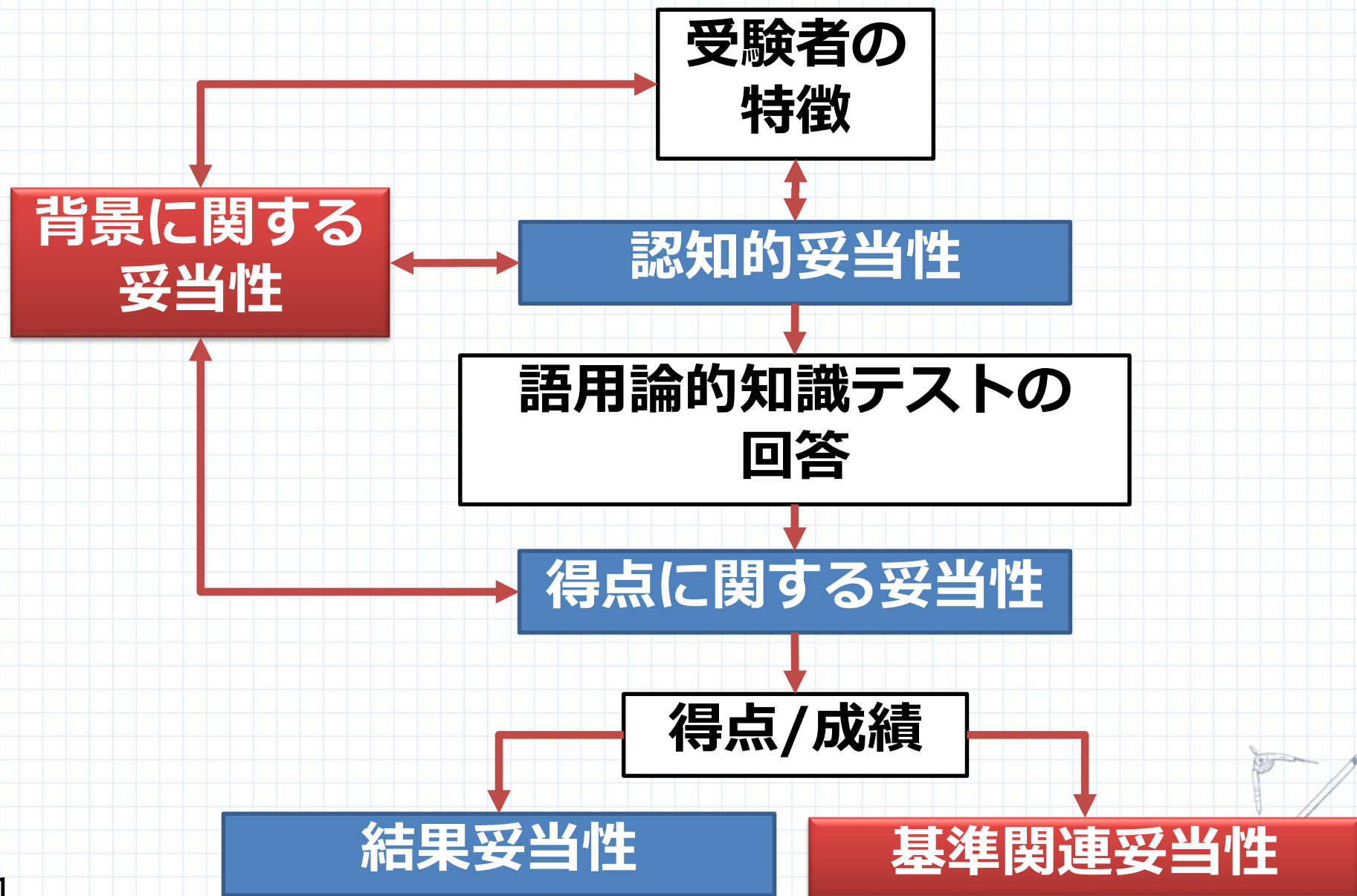


尺度	「相互行為への関与」の記述
3	・ インタラクション中ずっと、理解を共有できる。 ・ 相互行為への関与を示す証拠がある。 例：質問の明確化、あいづち、承認など
2	・ 理解を共有できないことがある。 ・ 相互行為への関与を示す証拠がいくつか見られるが、一貫していない。
1	・ 談話標識がない ・ 理解の共有に至らない



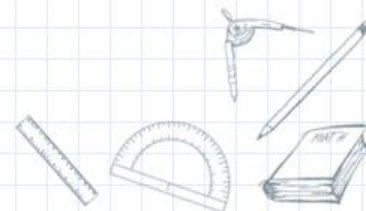
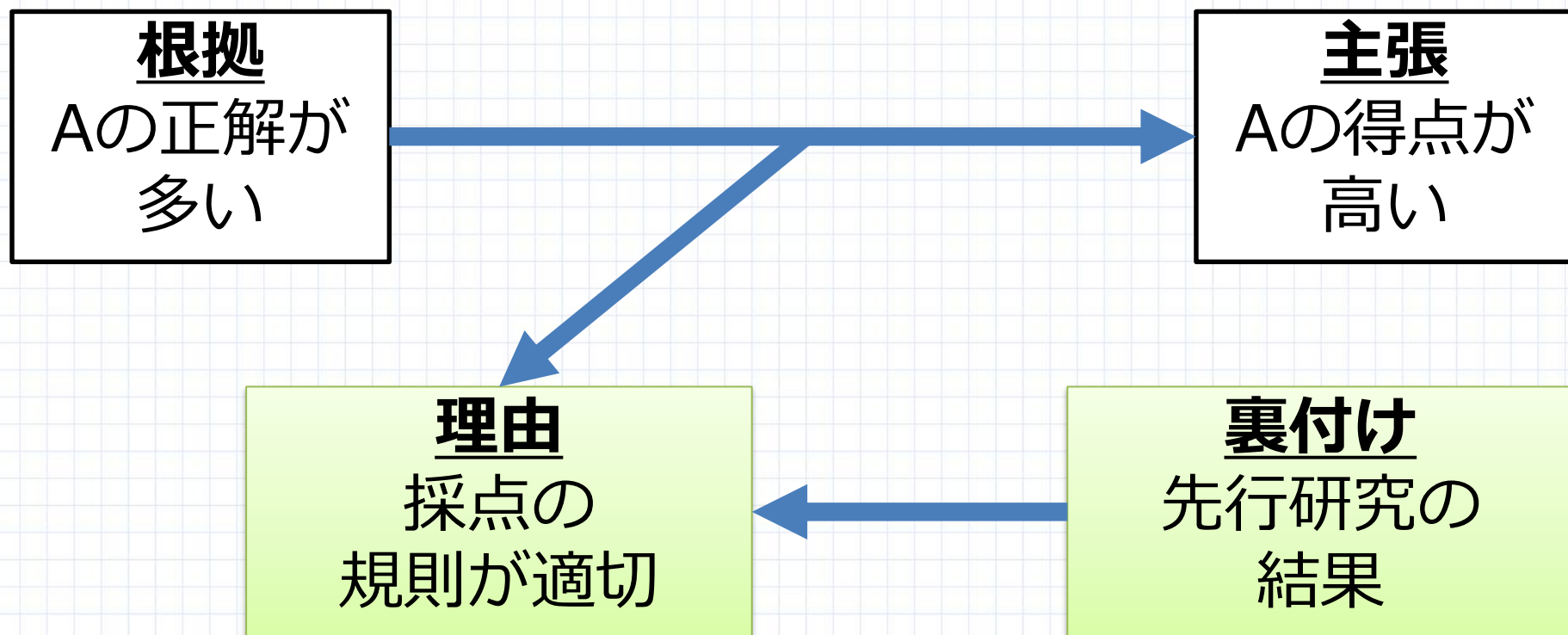


語用論的知識テスト(糸満, 2011) の検証



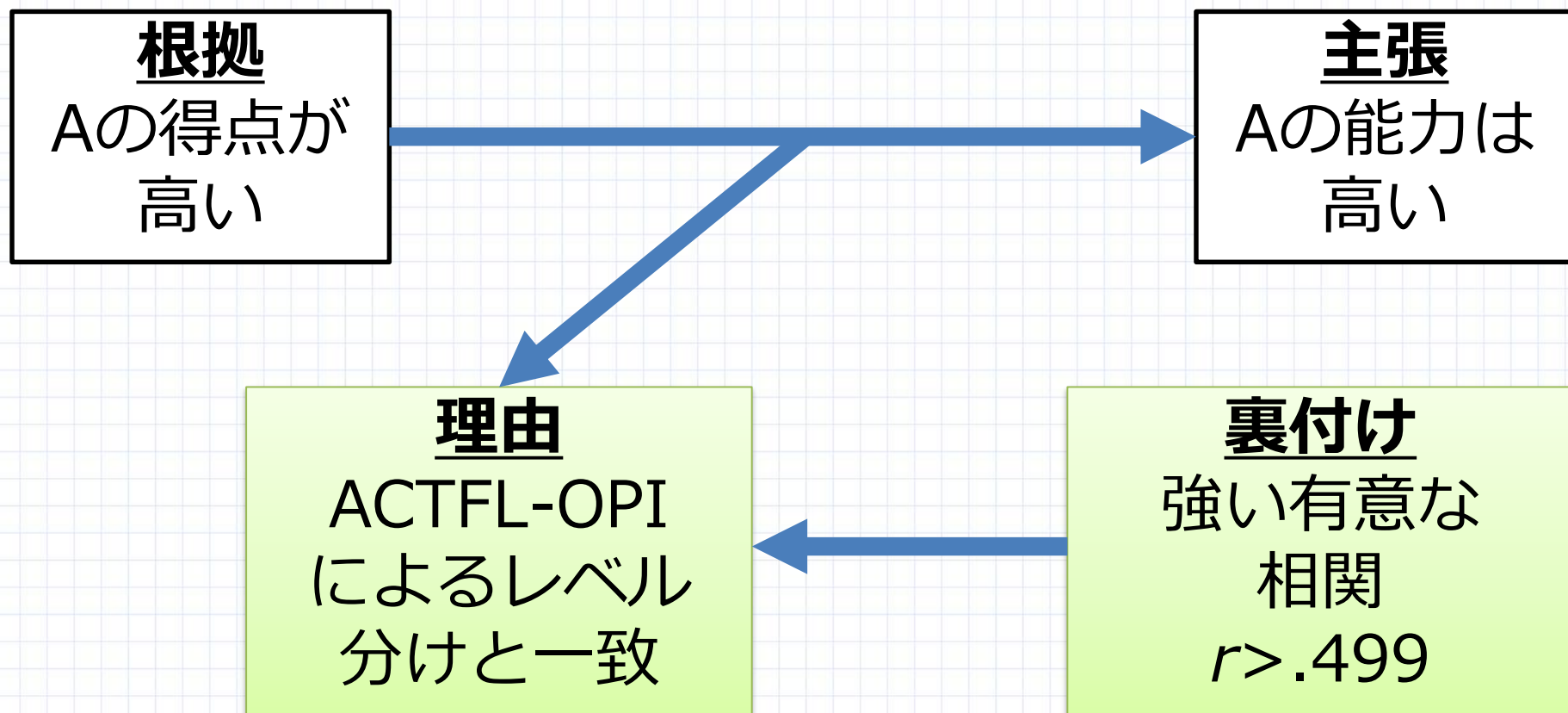


背景に関する妥当性の検討





基準関連妥当性の検討



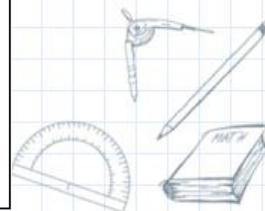


外国人児童生徒のためのJSL対話型 アセスメントDLA (文部科学省)



- 相互行為能力の重要性をふまえた評価法（ダイナミック・アセスメント）の一形態
- テスターが子どもと一対一で実施
- <はじめの一步>（「導入会話」と「語彙力チェック」）、

<話す>、<読む>、
<書く>、<聴く>

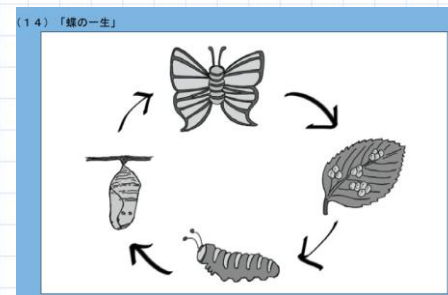




DLA <話す> の妥当性検証 (森重・渡部, 2016)



- インタビュー、ロールプレイ、絵描写：
基礎タスク、対話タスク、認知タスク
- ルーブリック：6段階尺度
＋記述（話の内容・まとめり、
文・段落の質、文法的正確度、
語彙、発音・流暢度、話す態度）
- 小学3年生15名
- テスター（採点者） 2名





ループリックの一部

(文部科学省)

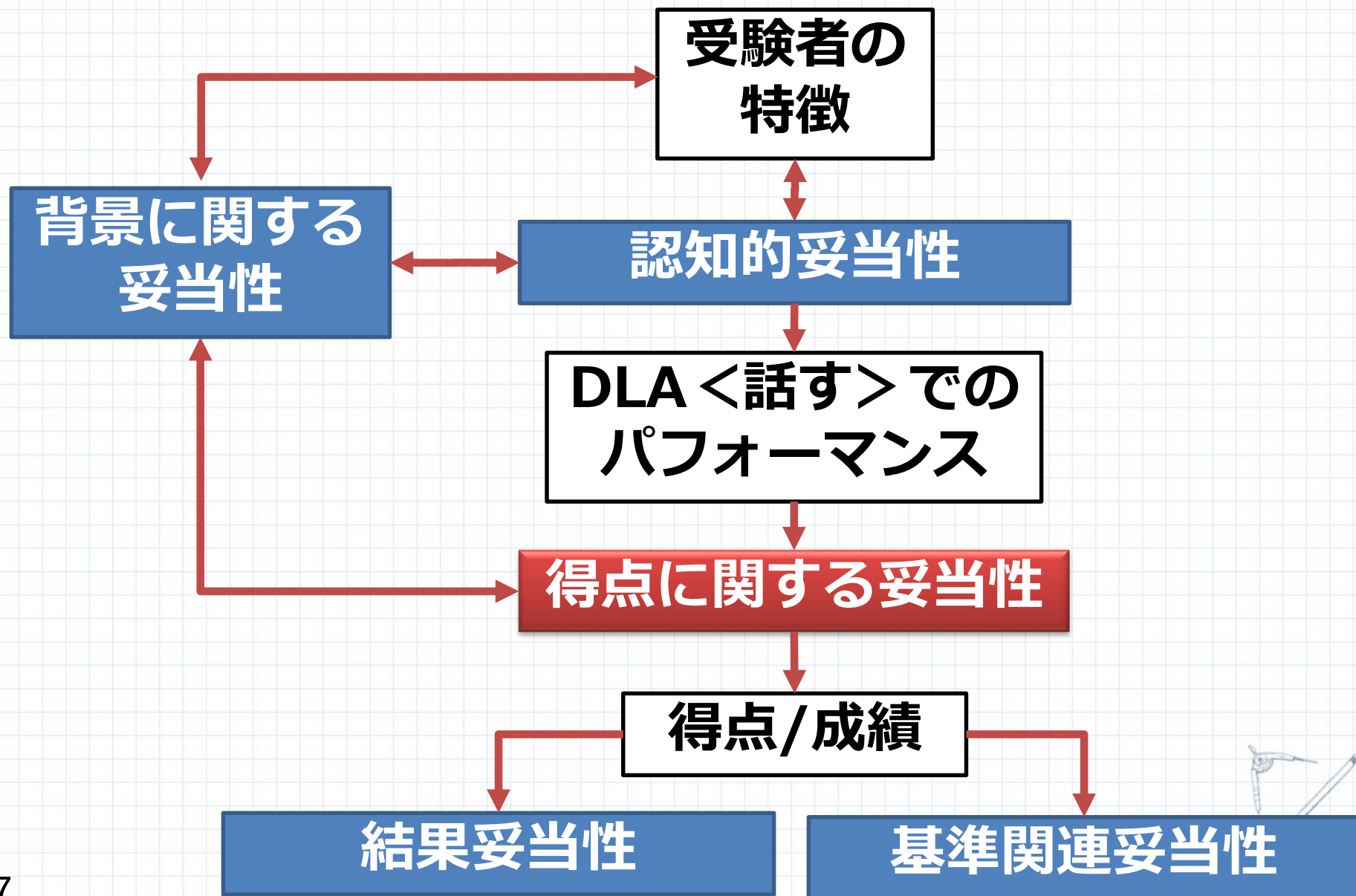


ステージ (尺度)	話の内容 まとめり	文・段落の質	文法的正確度
6	年齢相応の教科内容と関連した認知タスクがこなせる	まとまった話が1人でできる	文法的正確度が高い
4	対話タスクがこなせる	文を生成し、ある程度連文ができる	連文レベルで誤用がほとんど目立たない
1	基礎タスクの質問にいくつか答えられる	一語文	単語レベル



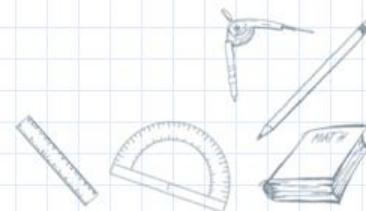
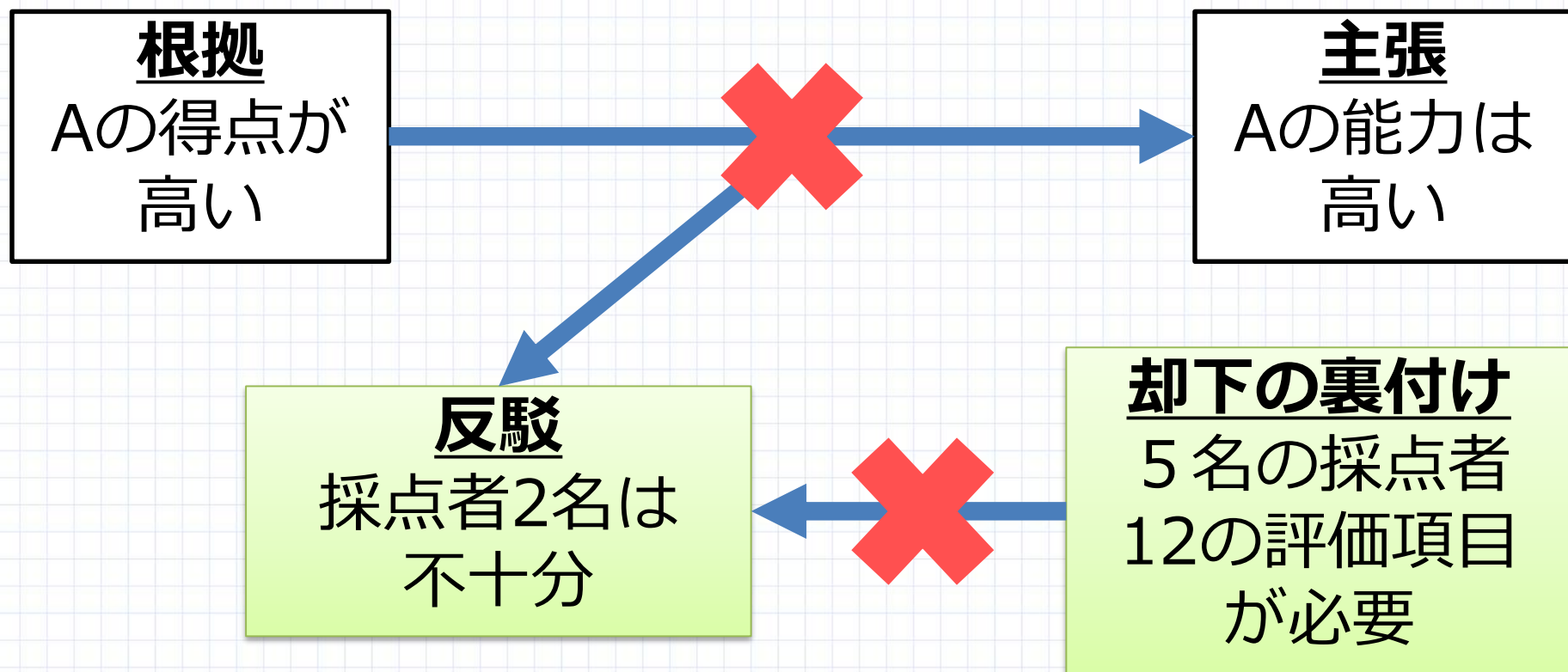


DLA <話す> の妥当性検証 (森重・渡部, 2016)



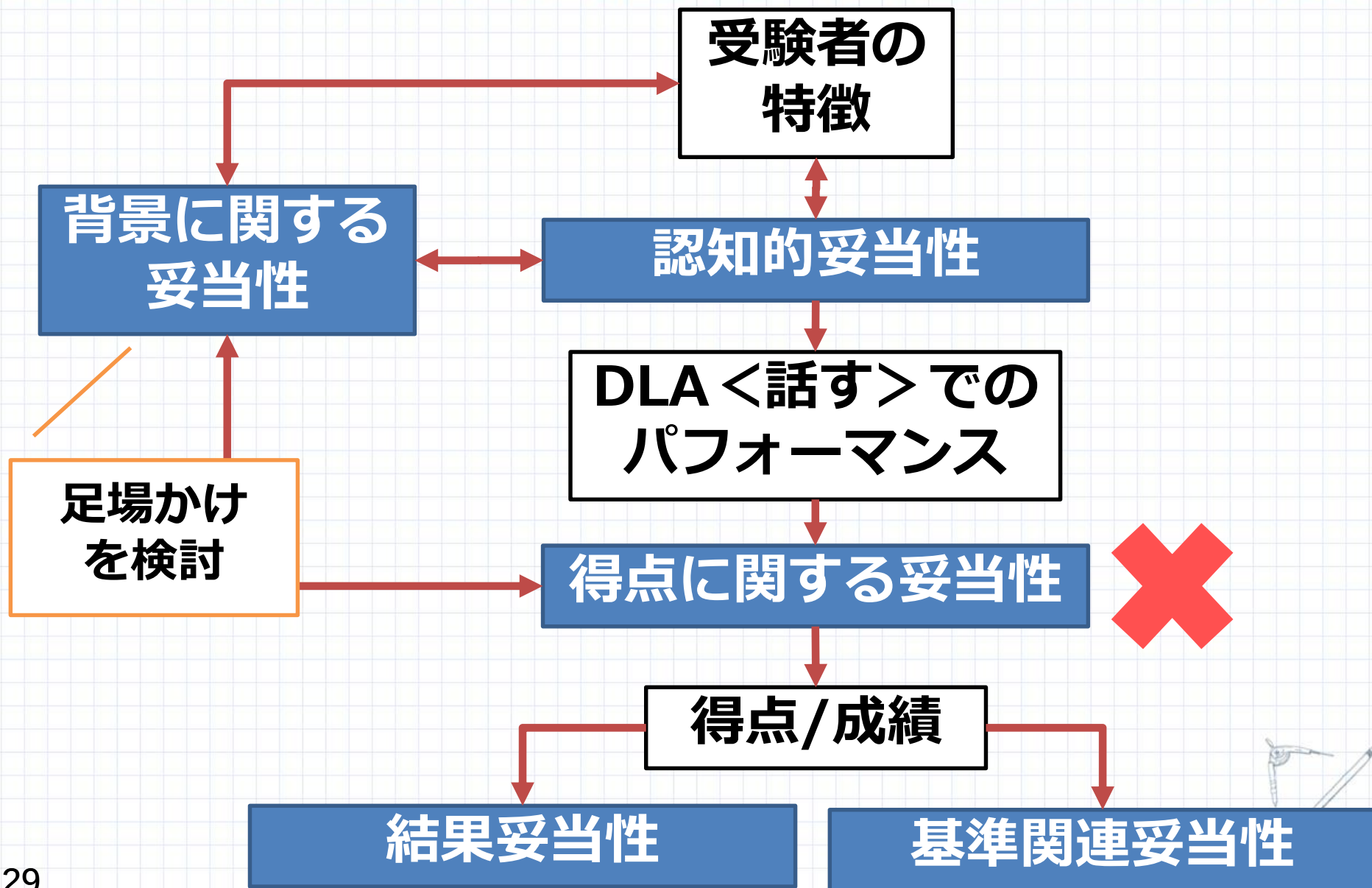


DLA <話す> の妥当性検証 (森重・渡部, 2016)



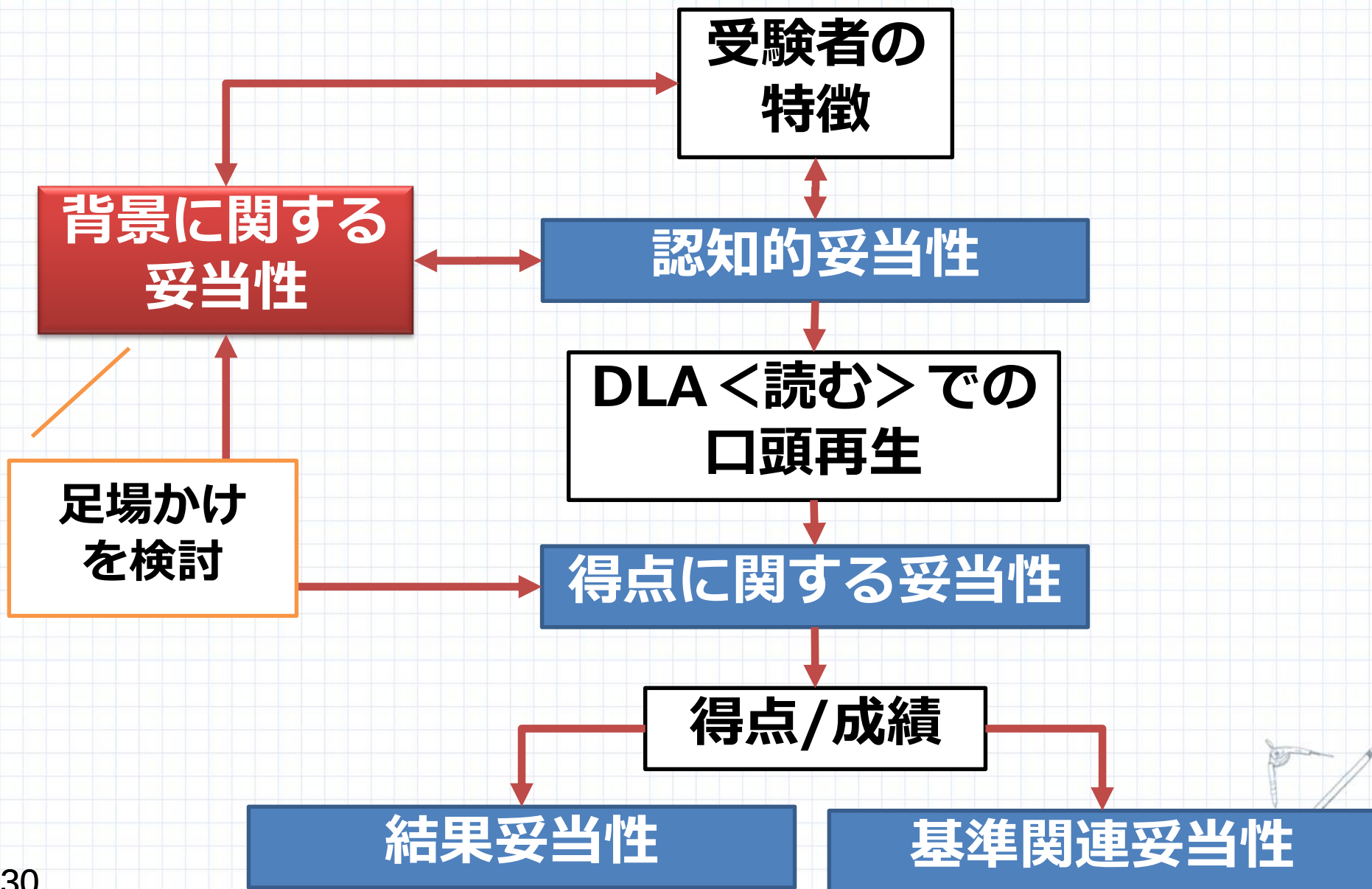


DLA <話す> の妥当性検証 (森重・渡部, 2016)





DLA <読む> の妥当性検証 (永田・渡部, 印刷中)

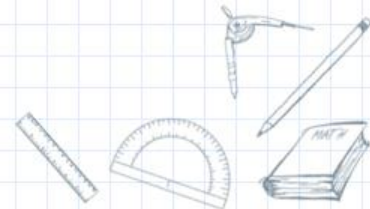




DLA <読む> の実践ガイドにおける 口頭再生の足場かけ



1. 途中で児童の読みや話を遮ったり，否定したり，訂正したりしてはいけない。
2. あらすじ・要旨の口頭再生ができない場合でも，
「一緒に最初から思い出してみよう」
「初めに誰がでてきましたか」
「それから？」
「最後にどうしましたか」
等と声かけをし，児童の発話を最大限に引き出す。

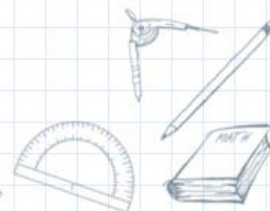




DLA <読む> の妥当性検証 (永田・渡部, 印刷中)



- 音読、黙読、内容の予測、**読んだテキストのあらすじ・要旨の口頭再生**、インタビュー（テキストの内容・意見、読書習慣）
- 談話分析：口頭再生の足場かけを抽出
 - － あいづち
 - － 先取り
 - － パラフレーズ
 - － 問題点の明確化
 - － 発話促進
 - － 問い
- 小学6年生15名、テスター3名





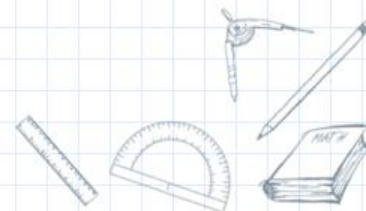
パラフレーズと発話促進の例



受験者A：（口頭再生、略）鉄腕アトムしかわからん。

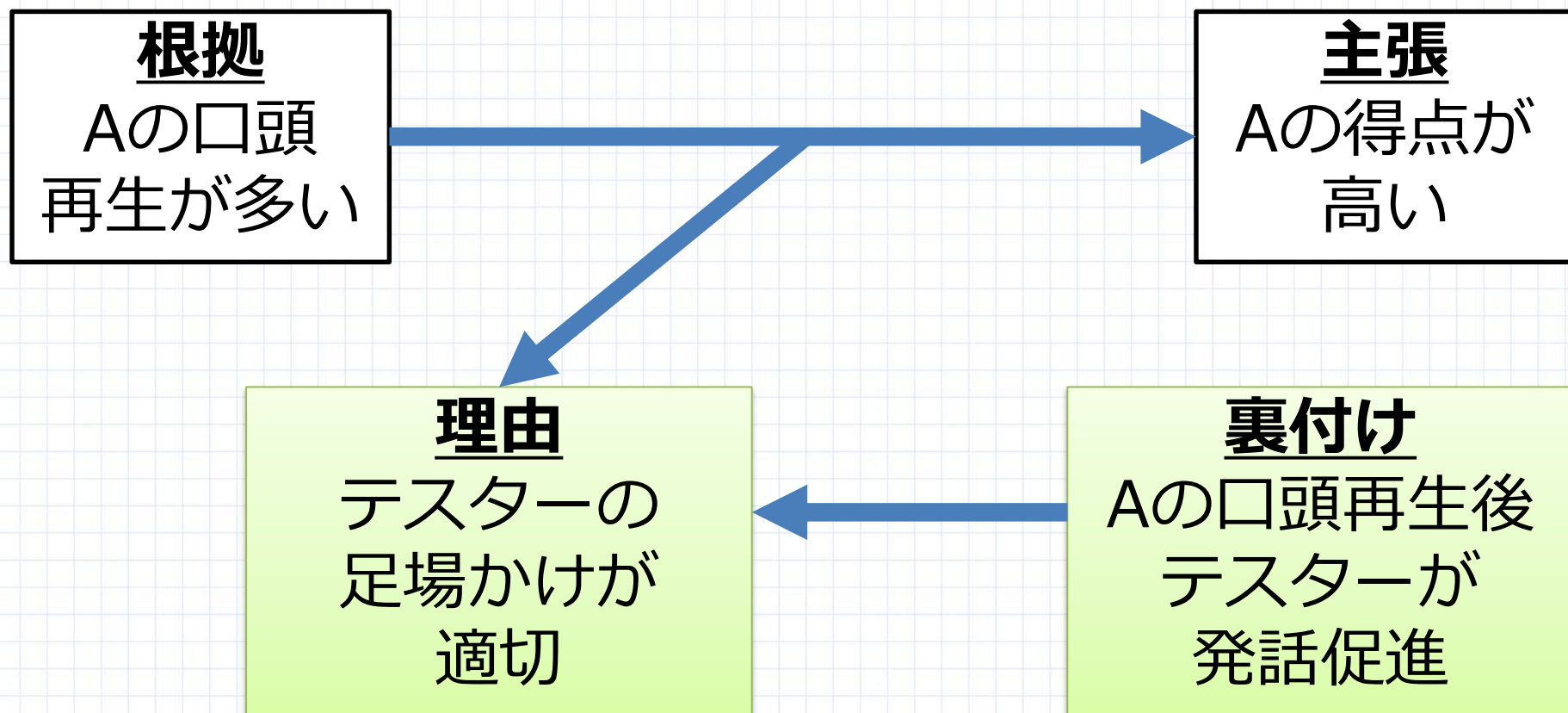
テスター：鉄腕アトムしかわからん？
（笑）【パラフレーズ】

はい。じゃあちょっとおも，思い出して
みようか。【発話促進】うん。でもすごいすごい。



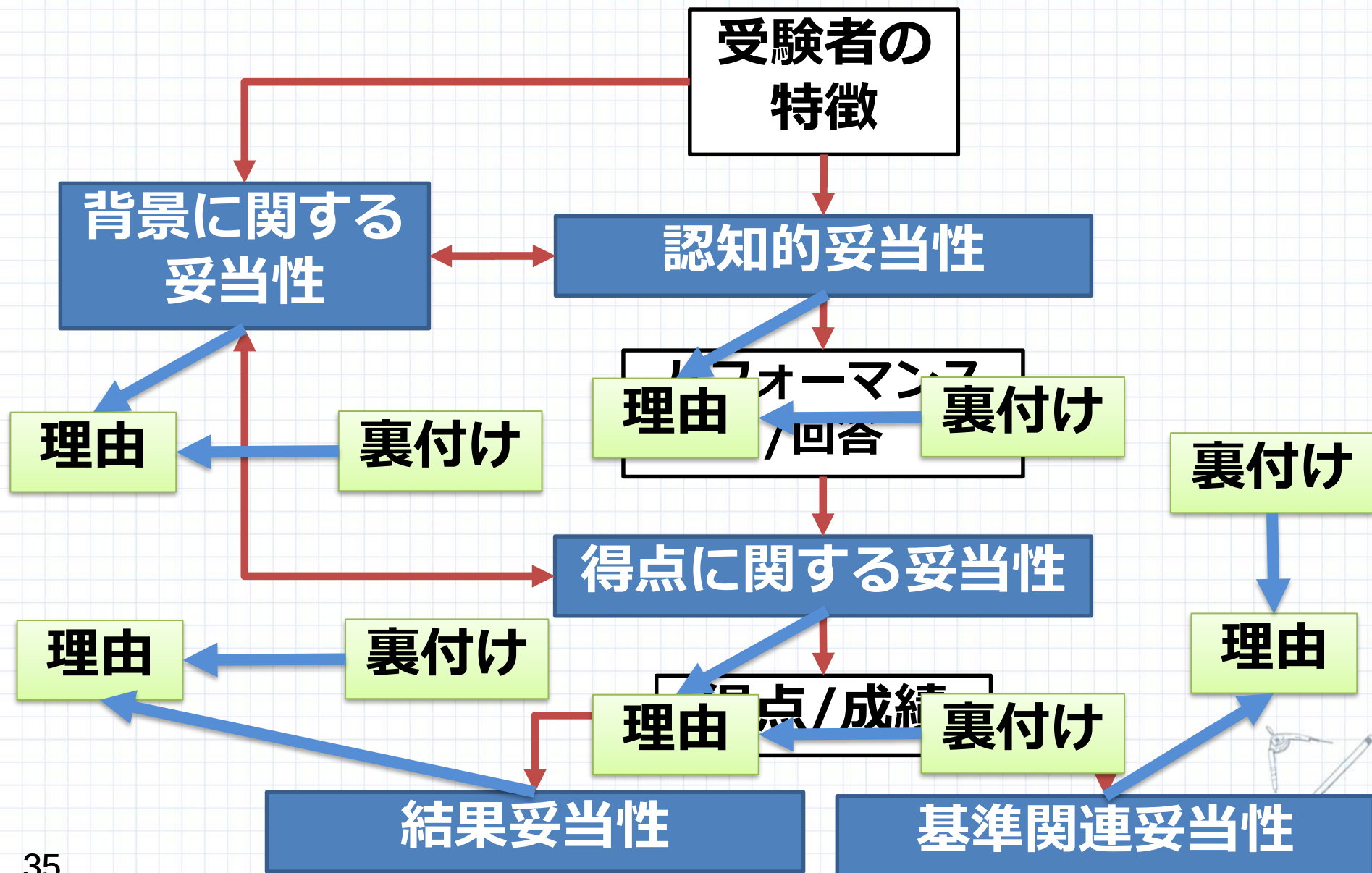


背景に関する妥当性の検討



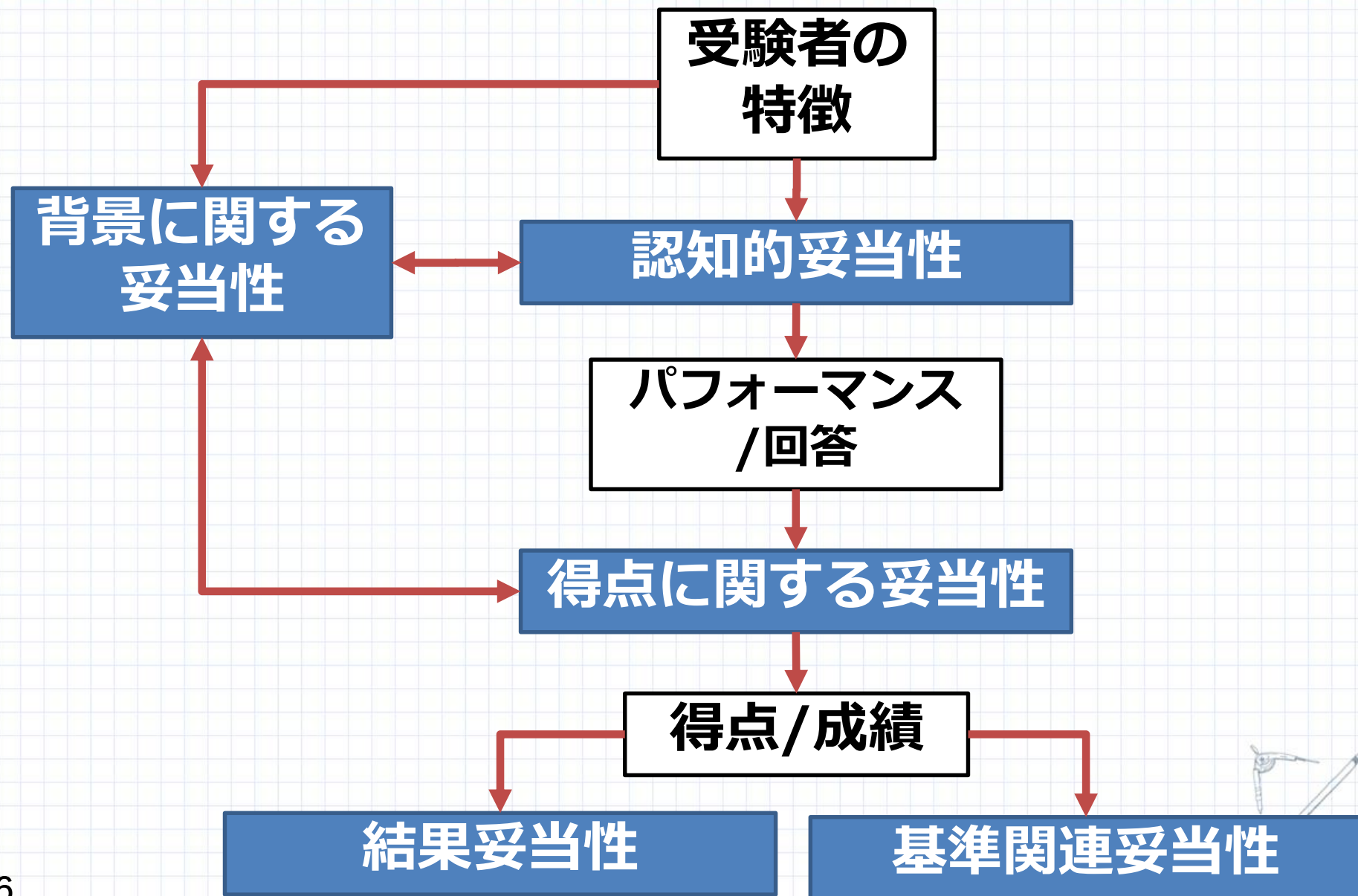


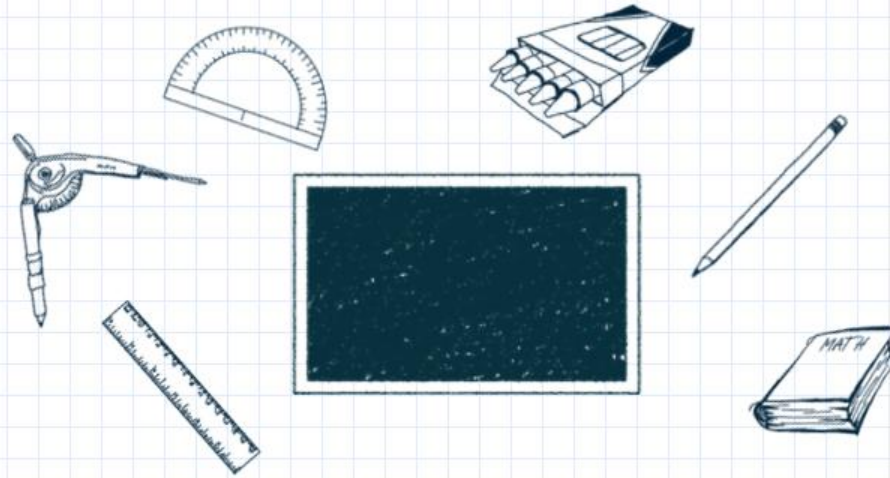
口頭運用能力の妥当性の検証





口頭運用能力の妥当性の検証





語用論的指導・会話分析的アプローチ による会話授業に必要な評価は？

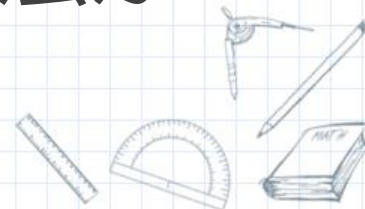
1. 実践と評価は背中合わせ。
2. 実践の目的は評価の目的。



語用論的指導における評価 (清水, 2016)



1. 個別具体的な文脈の中でコミニカティブな運用ができる。
2. 意図の伝達と相手への配慮のバランスを意識できる。
3. 言語形式・機能・文脈要因の関係に気づける。
4. 語用論的規範や傾向, 発話行為の談話構造や丁寧度の調整の方法が分かる。



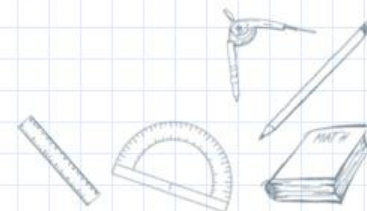


会話分析的アプローチにおける評価

(岩田, 2016)



1. 会話のやり方の知識（スピーチスタイル、連鎖構造、修復、決まり文句など）をリソースとして活用できる。
2. 会話の当事者が互いの言語・非言語的振舞いを検討することができる。
3. 会話の当事者が互いの関係の変化を検討することができる。



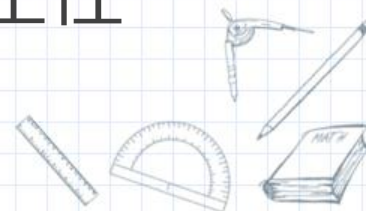


口頭運用能力の評価研究課題

(Nakatsuhara, 2012; Nakatsuhara & Field, 2012; Youn, 2015)



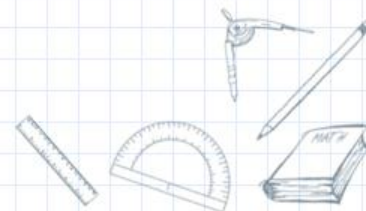
1. 多相ラッシュモデルによる評価者の
厳しさ、評価者間の一致度、必要な
タスクの数、評価基準の適切さ等
2. プロトコル分析による検討
 - インタビュアーが受験者にかかるリスニング
負荷（文法の複雑度、介入の数と長さ等）
 - 特定のレベルを判別する構成概念（マーカー
となる観測変数とは？）
 - ロールプレイやインタビューの真正性





3. 学習者による自己評価・ピア評価の妥当性

- 自己評価シートや内省の妥当性
- 心理尺度（異文化理解、ソーシャルスキル、対人関係意識、気づき等）の活用
- 形成的評価による学習の方向付け
- メタ語用論的側面や規範意識の変容

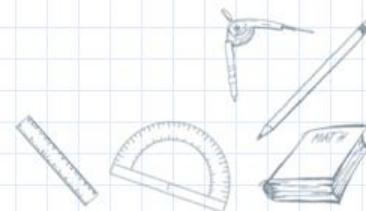


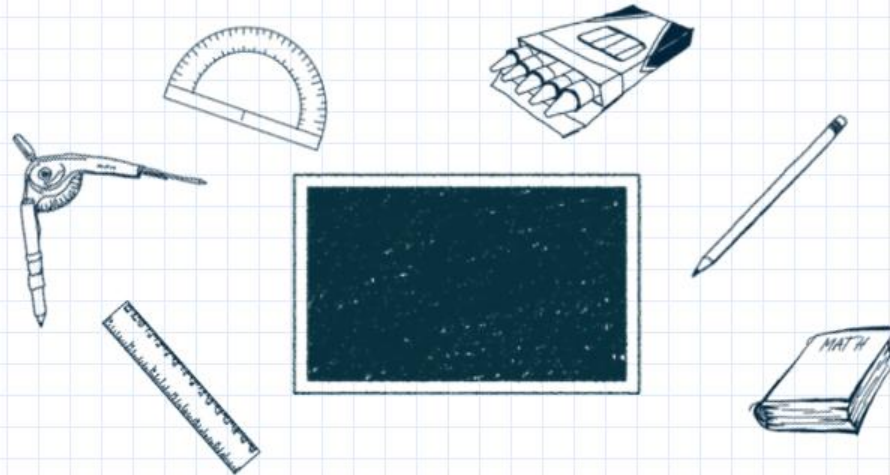


口頭運用能力の評価研究課題



4. テスター・採点者・学習者に対する
評価トレーニング（足場かけリスト
の提示、評価傾向チェック等）の効果
5. 会話分析的アプローチによる，テスト
が社会や人々に与える影響（結果
妥当性）の検証





口頭運用能力を どう評価すれば？

-
1. 実践・評価の目的に合った評価法を用いる。
 2. 実施前後に妥当性を検証する。
 3. 結果を共有する。

『第二言語としての日本語の習得研究』第20号
の投稿（一般/事例研究論文、展望論文）
締め切りは2017年1月31日です。

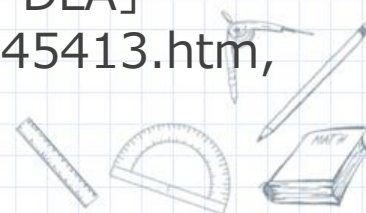
本研究は、2013年度JSPS科研費25284096
基盤研究（B）『アーティキュレーションを保証する言語能力ア
セスメント実施支援システムの構築』（研究代表者：渡部倫子）
の助成を受けた。



引用文献



- 石原紀子・アンドリュー D. コーエン (2015).『多文化理解の語学教育：語用論的指導への招待』東京：研究社
- 糸満真之 (2011).「日本語学習者の語用論的知識テストの開発：語用論的知識の測定はどこまで可能か」『言語教育評価研究』2, 26-34.
- 清水崇文 (2016).「コミュニケーションスキルが高まる授業を作る工夫－理論と実践をつないで－」『小出記念日本語教育研究会論文集』24: 102-109.
- 永田良太・渡部倫子 (投稿中).「外国人児童生徒のためのJSL対話型アセスメントDLA<読む>におけるテストの足場かけ」.
- 森重里保・渡部倫子 (2016).「一般化可能性理論を用いた外国人児童生徒のための JSL 対話型アセスメント DLA<話す>の検討」『日本語教育国際大会口頭発表予稿集』 http://bali-icjle2016.com/wp-content/uploads/gravity_forms/2-ec131d5d14e56b102d22ba31c4c20b9c/2016/08/morishigeriho_assessment-testing.pdf, 2016年10月20日閲覧.
- 文部科学省.「外国人児童生徒のためのJSL対話型アセスメントDLA」
http://www.mext.go.jp/a_menu/shotou/clarinet/003/1345413.htm, 2016年12月13日閲覧.





引用文献



- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010) *Language Assessment in Practice: Developing Language Assessments and Justifying Their Use in the Real World* .Oxford Applied Linguistics.
- Kane, M. T. (2006). *Validation*. *Educational measurement*, 4(2), 17-64.
- Nakatsuhara, F. (2012). *A research report on the development of the Test of English for Academic Purposes (TEAP) speaking paper for Japanese University entrants*. Project report submitted to STEP.
- Nakatsuhara, F., & Field, J. (2012). *A study of examiner interventions in relation to the listening demands they make on candidates in the GESE exams*. Project report submitted to Trinity.
- Rover, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, 28(4), 463-481.
- Weir, C. J. (2005a). *Language testing and validation*. Hampshire: Palgrave Macmillan.
- Youn, S. J. (2015). Validity argument for assessing L2 pragmatics in interaction using mixed methods. *Language Testing*, 32(2), 199-225.

