

Flucher, G. (2010). *Practical Language Testing*. London: Hodder Education.

Chapter 3 Activities

3.5 Practising your calculation Skills III

■ 40 人のクラスでライティングのテストを行ったと想定してください。

- ・ 2 種類のテストフォーム (Form1, Form2) を使用し、フォームの提示順は生徒間でカウンターバランスがとられました。
- ・ 20 人の生徒が Form1, 残りの 20 人が Form2 を受験しました。
- ・ Master と Non-Master を分ける基準点 (cut score) は 8 章で扱われる方法によって定められました。
- ・ そしてテストの結果、2 つのテストフォームは以下のように生徒を分類しました。

	Classification	Form 1		Total
Classification		Master	Non-master	
Form 2	Master	26 (A)	3 (B)	29 (A + B)
	Non-master	5 (C)	6 (D)	11 (C + D)
Total		31 (A + C)	9 (B + D)	40 (N)

■ この表を参考に、

【1】

(1) 一致係数 (agreement coefficient P_0) を計算しなさい。

$$P_0 = \frac{26 + 6}{40} = .80$$

(2) 偶然の一致率 (P_{chance}) を計算しなさい。

$$P_{chance} = \frac{29 \times 31 + 11 \times 9}{40^2} = .56$$

(3) カッパ係数を計算しなさい。

$$k = \frac{.8 - .56}{1 - .56} = .54$$

【2】

(1) これらの統計値からこのライティングの評価についてどんなことが分かりますか。

■ Form1 と Form2 の構成概念がおそらく同様であると考えると $.54 =$ 中程度の信頼性は決して高いものではない。このような 2 種類のフォームを比べる場合少なくとも $.80$ 程度の値が欲しい。

(2) これら2つのフォームを使用する前に、どんなことをしますか。

■ 信頼性をあげるために、(a) 項目自体の質を改善する、もしくは、(b) しかるべき基準に忠実に採点できるよう採点者をトレーニングする、という2つのアプローチが考えられる。エッセイなど項目数が少ないテストを評価する場合は、後者の方法が重要である。

3.6 Practising your calculation skills IV

■ あなたのクラスでプレイスメントテストを行ったと仮定します。

- ・テストは二者択一の (dichotomously scored) 100 項目からなり、1 問 1 点です。
- ・基準点は 70 点に定められました。
- ・平均正解率は .08 で、正解率標準偏差は .19 でした。

【1】

(1) $\Phi \lambda$ を計算しなさい。

$$\Phi \lambda = 1 - \frac{1}{100 - 1} \left(\frac{.8(1 - .8) - .19^2}{(.8 - .7)^2 + .19^2} \right) = .96$$

(2) 信頼区間 (CI) を計算しなさい。

$$CI = \sqrt{\frac{.8(1 - .8) - .19^2}{100 - 1}} = .017$$

3.7 It's good to talk!

■ 外部テストと、内部テスト (tests used in your own institutions; classroom assessment) の「リアル」な違いをあげてみましょう。

- ・それぞれの長所や短所は何か。
- ・教室内評価は単に「小規模なテストイング (small-scale testing)」とっていいのか。

☆ 外部テストの長所

- (1) しかるべき機関 (e.g., ETS) が作成したテストは妥当性が高い。
- (2) high stakes なテストで外部テストを用いれば、教師個人が社会的責任や労力を負わずに済む。

☆ 外部テストの短所

- (1) 個々のクラスや生徒に応じた測定を行うことは難しい。
- (2) 費用がかかり過ぎる。

☆ 内部テストの長所

- (1) 生徒の情意面や、他の生徒との関わり合いを見ることができるので、動的な評価が可能になる。
- (2) 個々に応じた評価が可能で、学習を促すことが期待できる。

☆ 内部テストの短所

- (1) どの科目でも毎回テストが課されるとなると、生徒の負担は非常に大きくなる。
- (2) 学校外の集団と相対的に比較することが難しい。