Methodological issues of functional assessment in sport
[what, how, when and why]

Franco M. Impellizzeri, PhD, FECSS
Research and Development Dept, Schulthess Clinic, FIFA Center of Excellence, Zurich, Switzerland
CeRiSM, University of Verona, Rovereto, Italy

Methodological Issues of Functional Assessment in Sport
[What, how, when and why]

Franco M. Impellizzeri, PhD, FECSS
Research and Development Dept, Schulthess Clinic, FIFA Center of Excellence, Zurich, Switzerland
CeRiSM, University of Verona, Rovereto, Italy

Outline

- [x] **Introduction.** Theoretical framework: functional assessment in research and practice.
- [ ] **Quality criteria** of measurement properties for functional assessments (validity, reliability, responsiveness, etc.)
- [ ] Collection and storage of test results (database)

Introduction - Framework

Methodological issues of functional assessment in sport

Settings
- RESEARCH
  - Mechanisms (basic)
  - Intervention efficacy
  - Determinants of performance
- FIELD
  - Control of interventions at group and **INDIVIDUAL** level
  - Training optimization (e.g. determination of intensity and loads)

Introduction - Framework

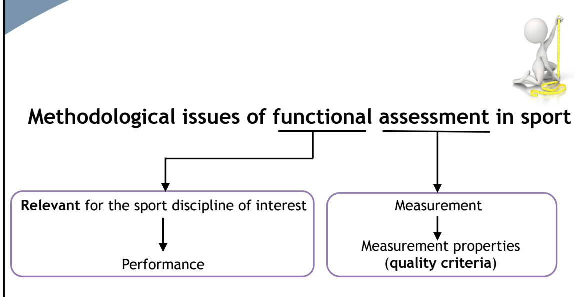
Methodological issues of functional assessment in sport

Introduction - Framework

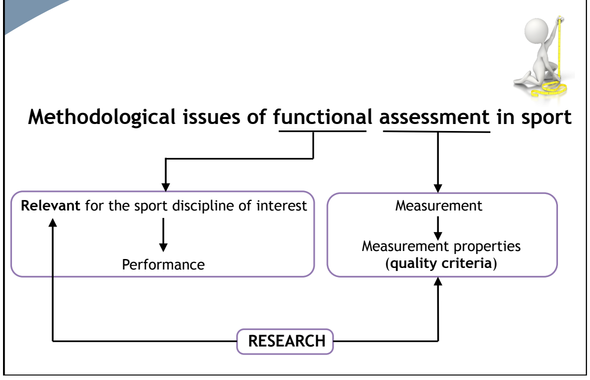
Methodological issues of functional assessment in sport

Measurement
↓
Measurement properties
(quality criteria)

Introduction - Framework



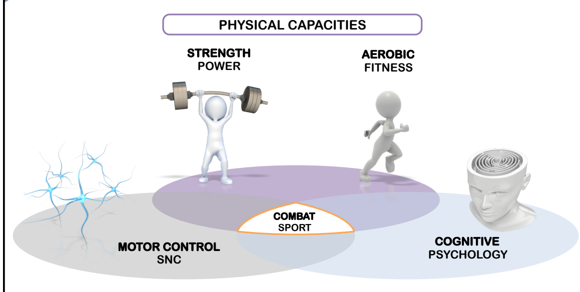
Introduction - Framework



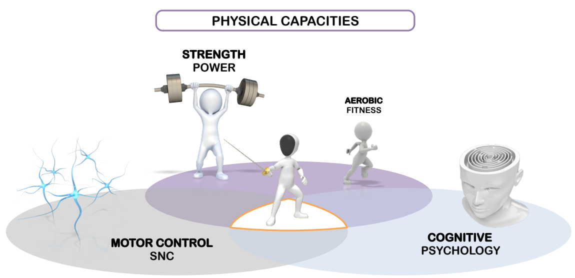
Summary – Take Home Message

1. The purpose of the research applied to sport (in the context of functional assessment) is to provide practitioners with valid measures (test)
2. … or to examine the validity of tests already used or available, always considering that field (coach experience) can give important and relevant suggestions for both training and testing
3. Sport scientists should not forget that coaches and athlete are the **stakeholder**

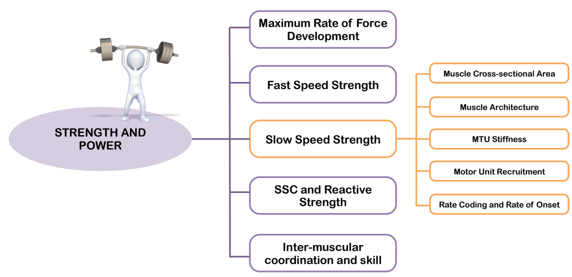
Introduction - Framework



Introduction - Framework

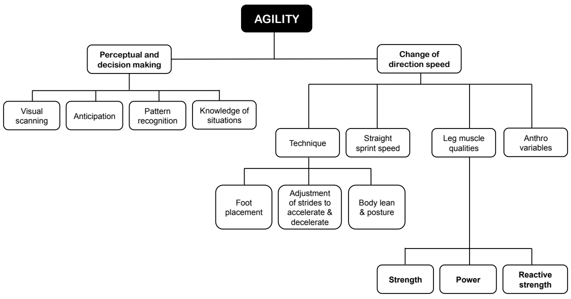


Introduction - Framework



[Cardinale et al. *Strength and Conditioning*. Wiley-Blackwell, 2011]

Introduction - Framework



[Cardinale et al. *Strength and Conditioning*. Wiley-Blackwell, 2011]

Introduction - Framework

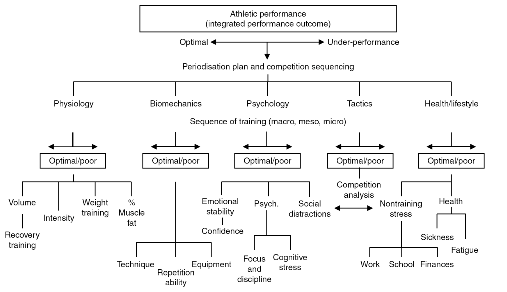


Fig. 5. A model of the contributing components to a measurable sport performance outcome called 'athletic performance'. **Psych** = psychological.

[**Smith.** Sports Med, 2003]

Introduction - Framework

GENERIC tests
Measures of single or few physical components (analytical measures)

Lower external validity

SPECIFIC tests
Integrated measures (i.e. integrating different determinants of the performance)

Higher external validity

❖ Same as training: the more specific usually the more externally valid (more related to performance)

Summary – Take Home Message

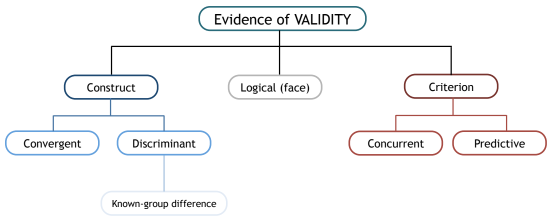
1. Functional tests measure ONLY ONE or SOME determinants of performance (low predicting ability)
2. We CANNOT pretend to predict performance from one or few tests!
3. Specific tests are more externally valid that generic (measures of generic prerequisites)

Outline

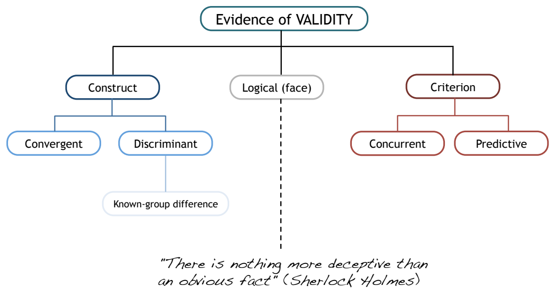
- ☑ **Introduction.** Theoretical framework: functional assessment in research and field.
- ☑ **Quality criteria** of measurement properties for functional assessments (validity, reliability, responsiveness, etc.)
- ☐ Collection and storage of test results (database)

Quality Criteria
Functional assessments

Measurement properties



Measurement properties



Measurement properties

VALIDITY

Validity is the degree to which the test measures what it purports to measure.

Validity is context specific (a test can be valid in one context but not in another)

[Ary et al, 2000]

Measurement properties

VALIDITY

Validity is the degree to which the test measures what it purports to measure.

Example of Aerobic Assessments:

1. VO2max
2. Lactate thresholds
3. Efficiency

} Valid measure of aerobic power and capacity: are they relevant measures for the performance and training diagnosis?

Measurement properties

Determinants of Endurance Performance

[Coyle, 1995; Joyner and Coyle, 2008]

1. **VO_{2max} (aerobic power)**
   - Determinant in heterogenous group
2. **Lactate thresholds**
   - Determinant in a group with homogenous aerobic power
3. **Efficiency**
   - Determinant in a group of endurance athlete with homogenous aerobic power and lactate threshold

Examples

Journal of Sports Sciences, 2005, **23**, 41–47

Taylor & Francis

Physiological correlates to off-road cycling performance

FRANCO M. IMPELLIZZERI,[1]* ERMANNO RAMPININI,[1] ALDO SASSI,[1] PIERO MOGNONI[2] and SAMUELE MARCORA[3]

[1]Human Performance Laboratory, S.S. MAPEI srl, Via Don Minzoni 34, 21053 Castellanza, Varese, Italy, [2]Consiglio Nazionale di Ricerca – Istituto Bioimmagini e Fisiologia Molecolare, Segrate, Milano, Italy and [3]School of Sport, Health and Exercise Sciences, University of Wales Bangor, George Building, Holyhead Road, Bangor, Gwynedd LL57 2PX, UK

Accepted 17 February 2004

Heterogeneous group (26 min time difference)

Table 3. Pearson correlations (r) between race time and physiological parameters expressed in absolute terms and scaled to body mass (n = 13)

	Race time *vs* absolute values		Race time *vs* values · BM^{-1}		Race time *vs* values · $BM^{-0.79}$	
	r	CI (95%) upper – lower	r	CI (95%) upper – lower	r	CI (95%) upper – lower
$\dot{V}O_{2peak}$	−0.66 *	(0.89 − 0.17)	−0.62 *	(0.87 − 0.10)	−0.68 *	(0.90 − 0.21)
Peak power output	−0.71**	(0.91 − 0.26)	−0.76**	(0.93 − 0.37)	−0.87 #	(0.96 − 0.62)
OBLA	−0.71**	(0.91 − 0.26)	−0.89 #	(0.97 − 0.67)	−0.94 #	(0.98 − 0.82)
Lactate threshold	−0.73**	(0.91 − 0.29)	−0.86 #	(0.96 − 0.59)	−0.90 #	(0.97 − 0.70)

Abbreviations: $\dot{V}O_{2peak}$, maximal oxygen uptake; OBLA, onset of blood lactate accumulation.
* $P < 0.05$; ** $P < 0.01$; # $P < 0.001$.

Examples

ORIGINAL ARTICLE

Correlations between physiological variables and performance in high level cross country off road cyclists

F M Impellizzeri, S M Marcora, E Rampinini, P Mognoni, A Sassi

Br J Sports Med 2005;39:747–751. doi: 10.1136/bjsm.2004.017236

Homogeneous group
(6 min time difference)

Table 2 Correlations (Pearson *r*) between race time and physiological variables expressed as absolutes and scaled to body mass values (n = 12)

	Race time *v* absolute values		Race time *v* values/BM[1]		Race time vs values/$BM^{0.79}$	
	r	95% CI	*r*	95% CI	*r*	95% CI
$\dot{V}O_{2max}$	−0.03	(−0.59 to 0.55)	−0.46	(−0.85 to 0.15)	−0.38	(−0.78 to 0.25)
$\dot{V}O_2$ at RCT	−0.18	(−0.68 to 0.44)	−0.66*	(−0.89 to −0.14)	−0.63*	(−0.88 to −0.06)
$\dot{V}O_2$ at VT	−0.01	(−0.58 to 0.57)	−0.38	(−0.78 to 0.25)	−0.30	(−0.75 to 0.33)
PPO	−0.11	(−0.64 to 0.50)	−0.48	(−0.83 to 0.13)	−0.43	(−0.81 to 0.19)
PO at RCT	−0.22	(−0.70 to 0.41)	−0.63*	(−0.88 to −0.06)	−0.61*	(−0.88 to −0.06)
PO at VT	−0.17	(−0.68 to 0.45)	−0.37	(−0.78 to 0.26)	−0.34	(−0.76 to 0.29)

BM, body mass; CI, confidence interval; PO, power output; PPO, peak power output; RCT, respiratory compensation point; $\dot{V}O_{2max}$, maximum oxygen uptake, VT, ventilatory threshold.
*p<0.05.

Examples

ORIGINAL ARTICLE

Correlations between physiological variables and performance in high level cross country off road cyclists

F M Impellizzeri, S M Marcora, E Rampinini, P Mognoni, A Sassi

Br J Sports Med 2005;39:747–751. doi: 10.1136/bjsm.2004.017236

Homogeneous group
(6 min time difference)

Table 2 Correlations (Pearson *r*) between race time and physiological variables expressed as absolutes and scaled to body mass values (n = 12)

	Race time *v* absolute values		Race time *v* values/BM[1]		Race time vs values/$BM^{0.79}$	
	r	95% CI	*r*	95% CI	*r*	95% CI
$\dot{V}O_{2max}$	−0.03	(−0.59 to 0.55)	−0.46	(−0.85 to 0.15)	−0.38	(−0.78 to 0.25)
$\dot{V}O_2$ at RCT	−0.18	(−0.68 to 0.44)	−0.66*	(−0.89 to −0.14)	−0.63*	(−0.88 to −0.06)
$\dot{V}O_2$ at VT	−0.01	(−0.58 to 0.57)	−0.38	(−0.78 to 0.25)	−0.30	(−0.75 to 0.33)
PPO	−0.11	(−0.64 to 0.50)	−0.48	(−0.83 to 0.13)	−0.43	(−0.81 to 0.19)
PO at RCT	−0.22	(−0.70 to 0.41)	−0.63*	(−0.88 to −0.06)	−0.61*	(−0.88 to −0.06)
PO at VT	−0.17	(−0.68 to 0.45)	−0.37	(−0.78 to 0.26)	−0.34	(−0.76 to 0.29)

BM, body mass; CI, confidence interval; PO, power output; PPO, peak power output; RCT, respiratory compensation point; $\dot{V}O_{2max}$, maximum oxygen uptake, VT, ventilatory threshold.
*p<0.05.

(no relation with gross and delta efficiency; unpublished results)

Measurement properties

Determinants of Endurance Performance

1. VO_{2max}, lactate thresholds and efficiency can be valid assessments but ONLY explain a small percentage of endurance performance in an homogeneous group of athletes (in MTB as in other sports).
2. More research is needed to identify the determinants of endurance performance and therefore the corresponding tests [e.g. psychobiological model; Marcora, 2009; 2010]

Measurement properties

VALIDITY

Validity is the degree to which the test measures what it purports to measure.

Validity is context specific (a test can be valid in one context but not in another)

[Ary et al, 2000]

Examples

Aerobic Fitness Variables Do Not Predict the Professional Career of Young Cyclists

PAOLO MENASPÀ[1], ALDO SASSI[1], and FRANCO M. IMPELLIZZERI[2,3]

[1]Mapei Sport, Castellanza, ITALY; [2]Department of Research and Development, Schulthess Klinik, Zurich, SWITZERLAND; and [3]Research Centre for Bioengineering and Motor Sciences, Rovereto, ITALY

Examples

TABLE 1. Comparison and discriminative ability of the anthropometric and physiological parameters of aerobic fitness between junior cyclists selected for the national team (NAT) and those not selected (non-NAT).

	NAT (*n* = 72)		Non-NAT (*n* = 237)				ROC	
	Mean	SD	Mean	SD	*P*	Partial η^2	AUC	95% CI
Age (yr)	17.5	0.5	17.5	0.5	0.639	0.001		
Height (cm)	179	7	177	6	0.025	0.016	0.576	0.499–0.653*
Body mass (kg)	70.0	8.0	65.0	6.4	<0.001	0.087	0.687	0.613–0.762†
% body fat[a]	5.6	2.0	5.9	1.9	0.352	0.003	0.454	0.375–0.534
$\dot{V}O_{2peak}$ ($L \cdot min^{-1}$)	5.017	0.622	4.584	0.564	<0.001	0.092	0.689	0.619–0.759†
$\dot{V}O_{2peak}$ ($mL \cdot kg^{-1} \cdot min^{-1}$)	72.1	7.4	70.7	6.7	0.130	0.007	0.557	0.478–0.635
$\dot{V}O_2$ at RCP ($L \cdot min^{-1}$)	4.321	0.530	3.891	0.498	<0.001	0.115	0.725	0.657–0.792†
$\dot{V}O_2$ at RCP ($mL \cdot kg^{-1} \cdot min^{-1}$)	62.1	6.8	59.9	6.1	0.013	0.020	0.594	0.517–0.670*
$\dot{V}O_2$ at V_T ($L \cdot min^{-1}$)	3.379	0.484	3.065	0.419	<0.001	0.086	0.673	0.603–0.743†
$\dot{V}O_2$ at V_T ($mL \cdot kg^{-1} \cdot min^{-1}$)	48.5	6.1	47.3	5.4	0.096	0.009	0.552	0.475–0.630

[a] *n* = 67 for the NAT, *n* = 202 for the non-NAT.
* $0.01 < P < 0.05$.
† $P < 0.001$.

- Some parameters of aerobic power and capacity may be used for selecting young athletes with more potential to excel in their categories

Examples

TABLE 4. Results of the multimodal logistic regression.

	Categories	Parameters	B	SE	P	OR	95% CI Lower	95% CI Upper
Model 1	PFL	$\dot{V}O_{2peak}$ (L·min^{-1})	1.051	0.044	0.021	2.860	1.168	7.003
		$\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1})	0.005	0.457	0.901	1.005	0.923	1.096
		Non-NAT	−1.561	0.523	0.003	0.210	0.075	0.585
	PCL	$\dot{V}O_{2peak}$ (L·min^{-1})	0.187	0.769	0.612	0.677	0.150	3.057
		$\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1})	−0.391	0.063	0.003	1.206	1.066	1.364
		Non-NAT	−0.030	0.881	0.973	0.970	0.173	5.454
Model 2	PC	$\dot{V}O_{2peak}$ (L·min^{-1})	0.863	0.628	0.170	2.369	0.692	8.112
		$\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1})	−0.001	0.060	0.991	0.999	0.888	1.124
		Non-NAT	−0.172	0.754	0.820	0.842	0.192	3.692
	PT	$\dot{V}O_{2peak}$ (L·min^{-1})	−0.594	0.473	0.210	1.810	0.716	4.579
		$\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1})	−0.095	0.044	0.031	1.099	1.008	1.198
		Non-NAT	1.751	0.555	0.002	0.174	0.058	0.515

0 athletes were correctly classified as professional cyclists!

Examples

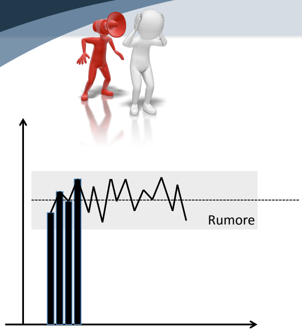
TABLE 4. Results of the multimodal logistic regression.

	Categories	Parameters	B	SE	P	OR	95% CI Lower	95% CI Upper
Model 1	PFL	$\dot{V}O_{2peak}$ (L·min^{-1})	1.051	0.044	0.021	2.860	1.168	7.003
		$\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1})	0.005	0.457	0.901	1.005	0.923	1.096
		Non-NAT	−1.561	0.523	0.003	0.210	0.075	0.585
	PCL	$\dot{V}O_{2peak}$ (L·min^{-1})	0.187	0.769	0.612	0.677	0.150	3.057
		$\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1})	−0.391	0.063	0.003	1.206	1.066	1.364
		Non-NAT	−0.030	0.881	0.973	0.970	0.173	5.454
Model 2	PC	$\dot{V}O_{2peak}$ (L·min^{-1})	0.863	0.628	0.170	2.369	0.692	8.112
		$\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1})	−0.001	0.060	0.991	0.999	0.888	1.124
		Non-NAT	−0.172	0.754	0.820	0.842	0.192	3.692
	PT	$\dot{V}O_{2peak}$ (L·min^{-1})	−0.594	0.473	0.210	1.810	0.716	4.579
		$\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1})	−0.095	0.044	0.031	1.099	1.008	1.198
		Non-NAT	1.751	0.555	0.002	0.174	0.058	0.515

0 athletes were correctly classified as professional cyclists!

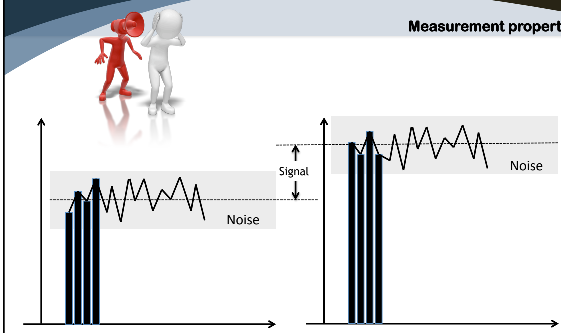
- … but they cannot be used for predicting “talent” (ability to compete at professional level);

Measurement properties



Noise <

Measurement properties



Noise < Signal

Measurement properties

- Reliability can be interpreted as the “noise” of a measure (instrumental errors, biological variability, etc.)

Absolute reliability (agreement)
Degree to which repeated measurements vary for individuals (longitudinal assessments)

Bland & Altman plot; SEM, Standard Error of Measurement

Relative reliability
Degree to which individuals maintain their position in a sample with repeated measurements (cross-sectional comparisons)

ICC, Intraclass Correlation Coefficient

[Atkinson and Nevill. Sports Med 1998; de Vet et al, J Clin Epidemiol, 2006]

Measurement properties

Minimal Detectable Change
From the SEM (or Typical Error; Hopkins, 2000) it is possible to determine the smallest change can be interpreted as real with an acceptable probability level.

Smallest Worthwhile Change
Smallest change relevant for the performance

[Atkinson and Nevill. Sports Med 1998; de Vet et al, J Clin Epidemiol, 2006; Hopkins et al., Sports Med, 2000]

Smallest Worthwhile Change

Minimal Detectable Change < **SIGNAL**

Training induced changes

❖ If the measurement error is the noise and the training effect (or the SWC) is the signal, responsiveness can be considered as the **signal:noise ratio**

[Atkinson and Nevill. Sports Med 1998; de Vet et al, J Clin Epidemiol, 2006; Hopkins et al., Sports Med, 2000; Amman et al, MSSE 2008]

Responsiveness (sensitivity to change)

Internal responsiveness
Ability of a measure to change over a particular prespecified time frame.

External responsiveness (longitudinal validity)
The extent to which changes in a measure over a specified time frame relate to corresponding changes in a reference measure.

[Husted et al., J Clin Epidemiol, 2000]

Yo Yo Intermittent Recovery Test

Reliability **NOISE** 5 to 8% (CV)	A 5% increase in YYIRT would indicate odds of about 5 to 1 the change is real.
Training-induced changes **SIGNAL** 9.9 to 30%	
Signal:Noise ratio Signal 2 to 6 times the noise	GOOD

(Bangsbo et al., 2008)

Repeated Sprint Ability test for soccer

	Reliability NOISE (SEM)	Training changes SIGNAL	Signal:noise ratio	Descriptor
RSE mean time	0.8%			
RSA best time	1.3%			
RSA decrement	30.2%			

(Impellizzeri et al., IJSM, 2008)

Repeated Sprint Ability test for soccer

	Reliability NOISE (SEM)	Training changes SIGNAL	Signal:noise ratio	Descriptor
RSE mean time	0.8%	2.1%	2.6	
RSA best time	1.3%	1%	0.8	
RSA decrement	30.2%	10%	0.3	

(Impellizzeri et al., IJSM, 2008)

Repeated Sprint Ability test for soccer

	Reliability NOISE (SEM)	Training changes SIGNAL	Signal:noise ratio	Descriptor
RSE mean time	0.8%	2.1%	2.6	GOOD
RSA best time	1.3%	1%	0.8	POOR
RSA decrement	30.2%	10%	0.3	VERY POOR

(Impellizzeri et al., IJSM, 2008)

Examples

Why the signal-to-noise is so important? An example

1. Reliability of time to exhaustion test = **20-25% CV**
2. Reliability of time trial = **1-3% CV**

Is Time-to-exhaustion reliable?

43

[Amman, Hopkins, Marcora, MSSE 2008]

Examples

Physical Fitness and Performance

Similar Sensitivity of Time to Exhaustion and Time-Trial Time to Changes in Endurance

MARKUS AMANN[1], WILLIAM G. HOPKINS[2], and SAMUELE M. MARCORA[3]

[1]The John Rankin Laboratory of Pulmonary Medicine, University of Wisconsin–Madison, Madison, WI; [2]Institute for Sport and Recreation Research, AUT University, Auckland, NEW ZEALAND; and [3]School of Sport, Health and Exercise Sciences, University of Wales–Bangor, Bangor, UNITED KINGDOM

ABSTRACT

AMANN, M., W. G. HOPKINS, and S. M. MARCORA. Similar Sensitivity of Time to Exhaustion and Time-Trial Time to Changes in Endurance. *Med. Sci. Sports Exerc.*, Vol. 40, No. 3, pp. 574–578, 2008. **Purpose:** There is widespread misunderstanding about the ability of constant-power tests to quantify changes in endurance performance. We have therefore compared the sensitivity of a constant-power test with that of a time trial for the effects of arterial oxygenation on endurance performance. **Methods:** Eight cyclists performed three constant-power rides to exhaustion and three 5-km time trials on a cycle ergometer in conditions of normoxia, hypoxia, and hyperoxia. After logarithmic transformation of performance times, sensitivity was calculated as the mean change in time divided by the error of measurement derived from the standard deviation of change scores. **Results:** In normoxia, performance times were 488 ± 77 s for the constant-power test and 454 ± 16 s (mean ± SD) for the time trial. The mean and standard deviation of the change in performance time from normoxia to hypoxia were much larger in the constant-power test (−45% ± 13%) than in the time trial (5.7% ± 1.6%); there was a similar disparity in the change from normoxia to hyperoxia (123% ± 37% and −4.1% ± 1.4%, respectively). However, sensitivity for the normoxia-hypoxia change in performance in the constant-power test (6.3, 90% confidence interval 4.3–11.4) was similar to that in the time trial (4.5, 3.0–8.2); sensitivities were also similar for the normoxia–hyperoxia changes (3.2, 2.1–6.0; 3.8, 2.5–6.9, respectively). *P* values for mean performance changes (range, 0.0002–0.000002) reflected these sensitivities. **Conclusions:** Time to exhaustion has sensitivity similar to that of time-trial time for the effects of arterial oxygenation and presumably other factors affecting endurance performance. Sensitivity need not be a concern when using constant-power tests to quantify changes in endurance performance. **Key Words:** CYCLING, EXERCISE TESTING, HYPEROXIA, HYPOXIA, RELIABILITY, ARTERIAL OXYGENATION

Examples

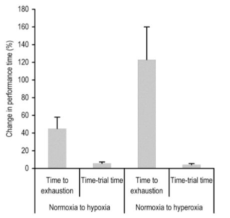


FIGURE 1—Percent changes in performance time during time trials and constant-power tests in hypoxia and hyperoxia compared with normoxia in eight male cyclists. Data are presented as means ± SD. The sign is reversed on the time to exhaustion in hypoxia and the time-trial time in hyperoxia.

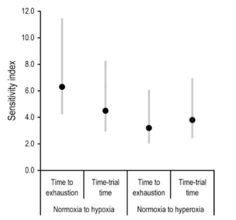


FIGURE 2—Sensitivity index for changes in performance time during time trials and constant-power tests in hypoxia and hyperoxia compared with normoxia in eight male cyclists. Data are presented as mean (*solid circle*) and 90% confidence interval (*line*). See Methods section for details on statistical calculations.

REVIEW ARTICLE

Sports Medicine 2001; 31 (3): 211-234
0112-1642/01/0003-0211/$22.00/0

Reliability of Power in Physical Performance Tests

Will G. Hopkins,[1] Elske J. Schabort[2] and John A. Hawley[3]

1 Department of Physiology, School of Medical Sciences, University of Otago, Dunedin, New Zealand
2 Sports Science Institute of South Africa, Cape Town, South Africa
3 Exercise Metabolism Group, School of Medical Sciences, RMIT University, Melbourne, Australia

Table III. Iso-inertial tests, sorted approximately by coefficient of variation (CV)

Reference	Participants; number of trials	Inter-trial time	Method of analysis	Movement; ergometer or instrument	Performance measure[a]	CV (%)	Change in mean (%)
Wiklander & Lysholm[34]	10 F, 10 M runners; 2	~1h?	?	Horizontal jump; tape	Distance or height (best of 2 jumps)[b]	1.8	?
				5-step jump; tape		1.5	
				Vertical jump; tape		4.3	
Risberg et al.[6]	12 F, 9 M elite athletes; 2	<1h?	Mean SD	Triple jump; ?	Distance[c]	2.1	?
				Vertical jump; ?	Height[c]	6.8	
		1-7wk	Mean SD	Triple jump; ?	Distance[c]	2.6	–0.6
				Vertical jump; ?	Height[c]	8.6	1.9
Viitasalo[35]	10 M athletes; 3-5	5-10 sec	?	Ball throw, low and high mass; photocells	Distance (best 2 of 3-5 throws)	2.5, 3.5	?
				Jump, unloaded and loaded; Ergojump	Height (best 2 of 3-5 jumps)	4.3, 6.0	
Ashley & Weiss[36]	50 F; 2;	2d	r & SD	Depth jump; Vertec	Height (best of 3 jumps)	3.1	?
				Restricted jump; Vertec		8.0	
Harman et al.[37]	18 M; 3	1-3 min	ANOVA	Vertical jump; AMTI LG6 force platform	Peak power[d]	3.3	?
					Height[d,e]	4.7	
Young et al.[38]	17 M athletes; 2	≥1d	?	Vertical jump, standing and striding; white board	Height (best of 3 jumps)	4.8, 3.6	–1.4, 3.5
				Vertical jump, standing and striding; Yardstick		3.8, 4.7	~0.0, 1.4
Bosco et al.[39]	4 F, 8 M throwers; 2	3d	SD diff	Squat lift, load = body mass; Ergopower	Mean power (best of 2 lifts)	3.7	0.8
		? min				5.0	5.4
Avis et al.[40]	53 M; 2	? min	r & SD	Leg press; weighted sled on rollers	Work[f]	5.1	1.2
					Peak power[f]	6.2	–0.9
Bosco et al.[41]	12 M boxers; 3 practice + 2	1 min	SD diff	Arm flexion, load = 5% body mass; MuscleLab-Bosco	Mean power	5.4[g]	–2.0[g]
Bassey & Short[42]	7 F + M; 2	~1wk	SD diff	Leg press; pedal + flywheel	Mean power, best of >4 reps	6.5	?

Measurement properties

Responsiveness (sensitivity to change)

What do you expect when your athlete improved in the tests?

External responsiveness (longitudinal validity)
The extent to which changes in a measure over a specified time frame relate to corresponding changes in a reference measure.

[Husted et al., J Clin Epidemiol, 2000]

Measurement properties

Responsiveness (longitudinal validity)

What do you expect when your athlete improved in the tests?

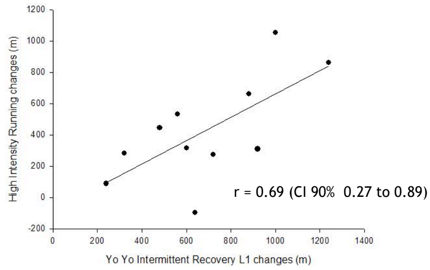
External responsiveness (longitudinal validity)
The extent to which changes in a measure over a specified time frame relate to corresponding changes in a reference measure.

Probably the most important test attribute!

[Husted et al., J Clin Epidemiol, 2000]

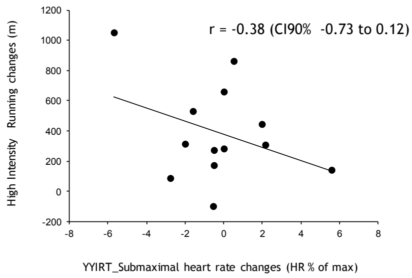
Responsiveness

Yo-Yo Intermittent Recovery Test Level 1



Responsiveness

Yo-Yo Intermittent Recovery Test SubMax



Summary – Take Home Message

Test validation process

Phase 1. Theoretical framework
Identification of the potential determinants of the performance.

Phase 2. Validity
Verification of the theoretical framework (relevance) and examination of test validity. When possible experimentally.

Phase 3. Reliability
Examination of reliability.

Phase 4. Responsiveness
Interpretation of the reliability in relation to the sensitivity to change and examination of the relation between change scores of the tests and the change in performance.

Phase 5. Interpretability
Assignment of qualitative meaning to a quantitative score (norm values, subgroup analysis, minimal detectable changes, smallest worthwhile changes…)

Outline

- ☑ **Introduction.** Theoretical framework: functional assessment in research and field.
- ☑ **Quality criteria** of measurement properties for functional assessments (validity, reliability, responsiveness, etc.)
- ☑ Collection and storage of test results (database)

Data Storing and Collection

A database is a collection of information that is organized so that it can easily be accessed, managed, and updated.

Database

Essential Database Characteristics:

1. Well-structured
2. Flexible
3. Data security
4. Data sharing
5. No redundancy
6. Feasible (data entering)
7. Quality control (e.g. plausible ranges of absolute and change scores)

Conclusion

Test selection process

Phase 1. Theoretical framework
Identification of the potential determinants of the performance.

Phase 2. Validity
Selection of the test demonstrated to be valid and relevant to the performance

Phase 3. Reliability
Select the tests for which the reliability is known (reliability may be population dependent!)

Phase 4. Responsiveness
Selection of the tests sensitive to changes.

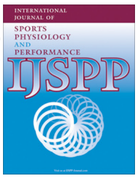
Phase 5. Interpretability
Selection of tests for which there are (or can be created) norm values, minimal detectable changes, smallest worthwhile changes…

Conclusion

TEST

INVITED COMMENTARY

International Journal of Sports Physiology and Performance, 2009, 4, 269-277



Test Validation in Sport Physiology: Lessons Learned From Clinimetrics

Franco M. Impellizzeri and Samuele M. Marcora

We propose that physiological and performance tests used in sport science research and professional practice should be developed following a rigorous validation process, as is done in other scientific fields, such as clinimetrics, an area of research that focuses on the quality of clinical measurement and uses methods derived from psychometrics. In this commentary, we briefly review some of the attributes that must be explored when validating a test: the conceptual model, validity, reliability, and responsiveness. Examples from the sport science literature are provided.

Keywords: physiological testing, measurement, validity, reliability, responsiveness, sport

"In conclusion, we believe that the application of more rigorous methods for the development and validation of physiological and performance tests would improve the quality of sport science research and professional practice.

… We hope that future investigations will assess all the relevant test attributes presented here, rather than leading to further **test proliferation**."

Olden Times

no.203.078 THE WORLDS OLDEST NEWSPAPER - since 1879 -

The new Omer Simpson's test !

The researcher and fitness coach Omer Simpson of the University of Springfield has developed a new test for soccer: the Omer Simpson's test!

This test is similar to those already existing, has not been appropriately validated, and we do not know how to interpret the results. The most original characteristics of the test is the name.

However, since this test is now used in professional teams, it can be considered valid.

The researcher and fitness coach Omer Simpson of the University of Springfield has developed a new test for soccer: the Omer Simpson's test!

This test is similar to those already existing, has not been appropriately validated, and we do not know how to interpret the results. The most original characteristics of the test is the name.

However, since this test is now used in professional teams, it can be considered valid.

"Quelli che s'innamoran di pratica sanza scienzia son come 'l nocchier ch'entra in navilio senza timone o bussola, che mai ha certezza dove si vada"

Leonardo Da Vinci

Thank you for the attention

Questions?