A psychophysical study of difficulty rating in rock climbing

Article *in* International Journal of Sport Psychology · January 1993

Delignières, D. et al. (1993). *International Journal of Sport Psychology*, *24*, 404-416.

A PSYCHOPHYSICAL STUDY OF DIFFICULTY RATING IN ROCK CLIMBING.

Didier Delignières[1], Jean-Pierre Famose[2], Chantalle Thépaut-Mathieu[3] & Philippe Fleurance[1]
(1) Laboratoire de Psychologie, INSEP, Paris.
(2) Université Paris-Sud, Orsay.
(3)Laboratoire de Biomécanique, INSEP, Paris.

Abstract: Because of the competitive evolution of rock-climbing, the rating scale of difficulty (RSD) has taken a central role in the organisation of this activity. However this scale is based on evaluative mechanisms, and little is known about its validity and metrical properties. Considering RSD as historically builded up with the accumulation of experts judgements, the assumption is proposed that RSD properties might be related to the individual processes of judgement. An experiment was conducted with the aim to test this hypothesis. 15 high-skilled climbers participated in the experiment. They had to perform 27 climbs on an artificial climbing wall. They were asked, after each path, to rate its difficulty, according to the RSD. Furthermore, subjects were asked to estimate, using the magnitude estimation method, difficulty, exertion, and required accuracy. Electromyographic data were collected on the right upper limb, with the aim of evaluating the objective exertion provided. Results showed that rating is more accurate as difficulty increases. RSD appeared as a logarithmic function of objective difficulty, and perceived difficulty as an exponential function of RSD. These results are discussed in relation with the psychological models of judgement.

Key-words: Rock-climbing, difficulty, rating scale, judgement.

Résumé: L'échelle de cotation en escalade a pris, du fait de l'évolution compétitive de la pratique, une importance centrale dans l'organisation de l'activité. Or, cette échelle repose sur des mécanismes d'évaluation, et l'on sait peu de chose sur sa valididé et ses propriétés métriques. Considérant que l'échelle de cotation a été historiquement élaborée à partir de l'accumulation des jugements d'experts, on peut penser que ses propriétés sont liées aux processus individuels de jugement. L'expérience présentée a pour but de tester cette hypothèse. 15 grimpeurs de haut-niveau participent à l'expérience, qui consiste à franchir 27 passages d'escalade, sur structure artificielle. On demande aux sujets, après chaque passage, d'en évaluer la difficulté selon l'échelle classique de cotation. Les sujets construisent en outre trois échelles, selon la méthode d'estimation des grandeurs, évaluant la difficulté, l'effort physique et la précision requise. Enfin, des relevés électromyographiques, effectués au niveau du membre supérieur droit, permettent de construire une échelle objective de l'effort requis. Les résultats indiquent que les cotations sont d'autant plus précises que les voies sont difficiles. L'étude des relations entre échelles montre que la cotation apparaît comme une transformation logarithmique de la difficulté objective, et que l'on peut décrire une relation exponentielle entre la cotation et la difficulté perçue. Ces résultats sont discutés en regard des modèles psychologiques de jugement.

Mots-clés: Escalade, difficulté, échelle de cotation, jugement.

Rock climbing is certainly the sport which presents the most advanced rating system for difficulty. Some activities (i.e. gymnastics) use classifications in three or four categories, but in the case of climbing, the scale is far more accurate. The climbing rating scale of difficulty (RSD) used in France[1] is graduated in degrees, each one subdivided in three levels. The degree is expressed by a number. From the first to fifth degree, subdivisions are expressed by addition of a sign - or + (e.g. 4-, 4 and 4+). From the sixth one, subdivisions are expressed by letters (e.g. 6a, 6b and 6c). Furthermore, climbers could make their ratings more precise, from the sixth degree, by adding a "+" if necessary (e.g. 6a+). Theoretically this scale has no upper limit. Currently the highest rating is 8c in cliffs, and 8a in bouldering.

The role of this scale became more important and diversified with the recent sportive and institutional evolution of rock climbing. Originally, it only served to rate the difficulty of mountain climbs, but now it is used for example to describe the level of a competition, to express a climber's expertise or to announce the requirements of a competitive examination (sport teacher, climbing trainer, etc..). It is surprising that such institutional functions were committed to an evaluative process which validity was not proved except empirically by a relative internal stability (Tiberghien, 1984).

Difficulty rating in rock climbing can be tackled from the point of view of the psychological models of judgment. One may consider that RSD was historically builded up with the accumulation of the judgments of the experts. Then RSD properties might be related to the individual processes of judgement.

In a judgment situation, the subject have not an absolute and autonomous representation of the object to judge: This representation must be placed in a larger one, related to a class of situations (Gonzales, 1988). The rating assigned to a path by a climber is probably related to his skill and knowledge levels in the activity. One may suppose that climbers give their RSD judgements by comparing the actual perceived difficulty with the difficulty they have perceived during earlier ascents, in classical and well-known paths, whose ratings are commonly admitted. Then the hypothesis could be proposed, that as the subject approach the level of difficulty he use to climb (i.e. near his maximal limit), the rating become more accurate.

Furthermore, the rating must be understood through the functions it assumes. A judgement is an adaptative response: it represents the subject's comprehension of the situation, and its position in this situation (Payne, 1982). During his sportive experience, the climber builds up a representation of the continuum of climbing difficulties. The rating locate the path on this continuum, and preserve its coherence. This continuum is individually confined by one's representation of its own maximal limit. One may suppose that near this limit appears a ceiling effect, and that more generally the relationship between objective difficulty and RSD must be negatively accelerated.

If this hypothesis is valid, a dissociation might be observed, at the high levels of difficulty, between RSD and perceived difficulty. Earlier works on the perception of difficulty or exertion have shown positively accelerated relationships between stimuli and sensations (Borg, 1962, Delignières, 1990). Then a positively accelerated relationship between RSD and perceived difficulty might be described.

The presented experiment has been conducted to verified these hypotheses.

[1] Different scales are used in other countries, for example Great Britain, Germany or USA.

METHOD.

Subjects.

15 expert climbers (average age 26 y. 8 m., sd 6 y. 2 m.) participated in the experiment. Their average skill level (on cliffs) was 7b on sight (i.e. the level of difficulty systematically performed at the first trial) and 7c+ after practise.

Experimental device.

The experiment used a specific artificial climbing wall, made of a large plate of plywood (4 meters high and 3 meters wide), fixed on a rigid metal frame. This plate could pivot around its upper axis, so it was possible to give it a variable inclination, from 110 to 70 degrees. Climbing holds of various forms and sizes could be adjusted on this plate (Figure 1).

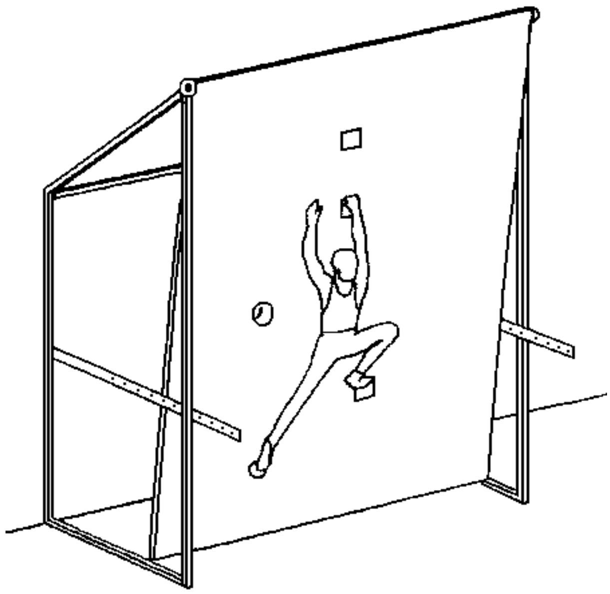

Figure 1: The experimental artificial climbing wall.

Traditionally, a distinction is made between athletic paths (i.e. requiring muscular force) and technical paths (i.e. requiring balance and accuracy; Tiberghien, 1984; Dupuy, 1987; Delignières, 1990). Since presently the only available way to obtain a valid measurement of objective difficulty in climbing tasks is to measure objective exertion, we have chosen athletic paths in our experiment.

The experimental tasks was made up of three holds, named H1, H2 and H3. Subjects had to grasp H3 with the left hand, using H2 with the right hand and H1 with the right foot. Three additional holds could help them, at the beginning of the movement, to balance themselves on the H1-H2 system. The localization of the holds does not vary between paths (Figure 2).

The dimensions of variation were the size (depth) of the holds H1 and H2, and the inclination of the wall. H1 and H2 were rectangular holds 10 centimeters wide. Their depth could be 1.2, 0.6 or 0.3 centimeters for H1, and 2.4, 1.2 or 0.6 centimeters for H2. Three inclinations of the plate were used (ß angle): 80, 95 and 110 degrees. Whatever the inclination, the useful surfaces of H1 and H2 were horizontal (Figure 3).

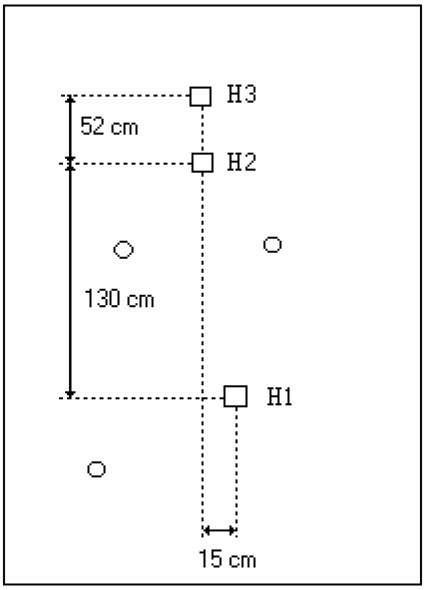

Figure 2: Localization of the holds on the wall (o: additional holds).

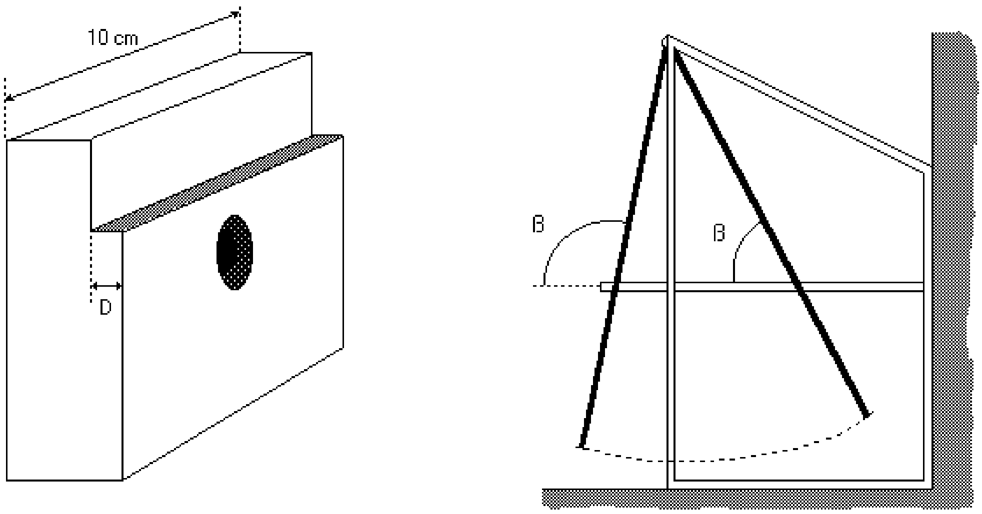

Figure 3: Up: experimental hold, and measurement of the depth (D); Down: measurement of the inclination (ß angle).

By a systematic combination of the three levels of inclination of the wall and of the three depths of each holds, we obtained 27 experimental paths. Depths and inclination have been empirically chosen to cover by their combination a wide range of the conceivable difficulties in rock climbing.

Table 1. Depth of the foot hold (H1, cms), depth of the hand hold (H2, cms), inclination of the wall (ß, degrees), average rating of difficulty (RSD) and number of successfull subjects, for each experimental path.

Task	H1	H2	ß	RSD	Subjects	Task	H1	H2	ß	RSD	Subjects
1	1.2	2.4	100	3	15	15	.3	1.2	85	5c+	15
2	.6	2.4	100	3+	15	16	1.2	.6	85	6a	15
3	.3	2.4	100	4	15	17	.6	.6	85	6b	15
4	1.2	1.2	100	3+	15	18	.3	.6	85	6b+	15
5	.6	1.2	100	4-	15	19	1.2	2.4	85	5a+	15
6	.3	1.2	100	4+	15	20	.6	2.4	85	5b+	15
7	1.2	.6	100	4-	15	21	.3	2.4	85	5c+	15
8	.6	.6	100	4+	15	22	1.2	1.2	85	6a	15
9	.3	.6	100	5b	15	23	.6	1.2	85	6b	15
10	1.2	2.4	85	4	15	24	.3	1.2	85	6b+	15
11	.6	2.4	85	4+	15	25	1.2	.6	85	6c+	10
12	.3	2.4	85	5a+	15	26	.6	.6	85	7a	7
13	1.2	1.2	85	5a	15	27	.3	.6	85	7b	6
14	.6	1.2	85	5b	15						

Procedure.

Because of their extreme difficulty, the tasks 17, 18, 23, 24, 25, 26 and 27 were always performed in this order, at the end of the experiment. The order of the others paths was systematically randomized between subjects. Subjects had for each task as many trials as is necessary. After practise and ascent of each path, they were asked to rate its difficulty, according to RSD.

Three other scaling were realized, using the method of magnitude estimation (Stevens, 1957, 1968/69). These estimations concern respectively:

- the difficulty of the path (the same criterion as that estimated with RSD);

- the physical exertion needed to perform the task;

- the required motor accuracy.

The two last criteria refered to the distinction athletic/technical, commonly admitted by climbers. The easiest task served as modulus, and received the mark 10 for the three scales. Subjects had to attribute marks to the other tasks in a ratio perspective: for example, a path twice as difficult as the modulus was marked 20 on the difficulty scale.

Furthermore, for 5 climbers of the experimental group, electromyographic data were collected on four muscles of the right upper limb, which is the most solicited in the hardest movement the 27 tasks (flexor digitorum profundis, palmaris, biceps bracchii and pectoralis major).

Data processing.

Because of the extreme difficulty of some paths and also of injuries suffered during the experiment, all subjects were not able to perform all the tasks. According to the numbers of paths performed and scored, subjects were distributed as follows:

27 paths........ 6 subjects
26 paths........ 1 subject
25 paths........ 3 subjects
24 paths........ 5 subjects

Two treatments were simultaneously performed, the first concerning the paths realized by all the subjects (24 paths, 15 subjects), and the second the subjects which had achieved the whole experiment (27 paths, 6 subjects).

With the aim of using statistical treatments such as correlations or regressions, the ratings obtained with the RSD were transformed into decimal numbers: 1 was attributed to the rating 1, 2 to the rating 1+, 3 to 2-, and so on... Then for each path the average score was computed (RSDsc). Then, these scores might be converted into ratings (RSD), by taking the nearest whole or intermediate corresponding rating (see Table 1).

The five subjects tested by electromyography had performed all the tasks. Electromyographic data were integrated. The given values were the maximal values observed for each muscle during one second. If the data within this standard time interval were too heterogeneous (e.g. in burst), this time interval was reduced to isolate the relevant data.. Then the average score was calculated, for each task and each muscular group.

Concerning the three ratio scales, the average score was calculated for each task.

RESULTS.

Ratings of difficulty (RSD).

Average ratings, computed from all the available data for each path, were distributed from 3 to 7b (Table 1). This represents a large sample of the conceivable difficulties in rock climbing. There was a negative correlation between difficulty and standard deviation (24 paths, 15 subjects: $r=-.742$, $p<.01$; 27 paths, 6 subjects: $r=-.595$, $p<.01$, Figure 4).

Difficulty, exertion and accuracy.

The three ratio scales elaborated during the experiment allowed us to study the respective contributions of exertion and accuracy in the building up of the global feeling of difficulty. A multiple regression analysis gave the following equations:

24 paths, 15 subjects: $Z_{PD} = (.927)Z_{PE} + (.071)Z_{PA} + .000$ $(r=.993)$

27 paths, 6 subjects: $Z_{PD} = (.831)Z_{PE} + (.170)Z_{PA} - .000$ $(r=.996)$

in which Z_{PD}, Z_{PE} and Z_{PA} respectively represented, in z-scores, the perceived difficulty, the perceived exertion and the perceived accuracy. With regard to the paths used in this experiment, perceived exertion seemed to have a more important load in the estimation of difficulty than accuracy.

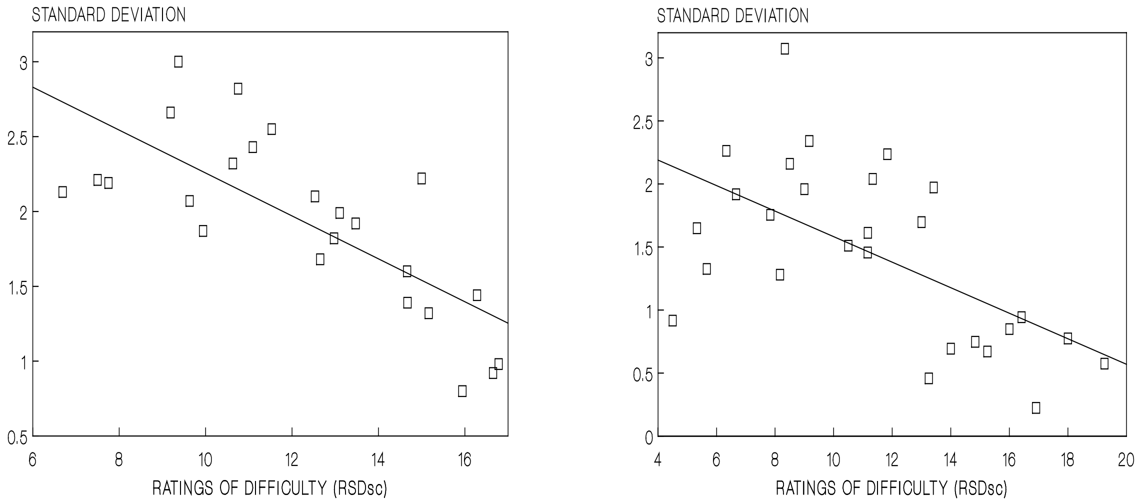

Figure 4: Transformed scores of difficulty (RSDsc): relationship between mean and standard deviation. Left: 24 tasks, 15 subjetcs; right: 27 tasks, 6 subjects.

Perceived exertion and objective exertion.

Rank correlations between perceived exertion and electromyographic data showed great differences between muscular groups:

-Flexor digitorum profondis	r=.989
-Palmaris	r=.937
-Biceps.brachii	r=.834
-Pectoral major	r=.558

The relevant stimulus for the building up of perceived exertion seems clearly located in the muscles of the anterior compartment of the forearm. Taking into account the load of exertion in the estimation of difficulty, one may consider that electromyographic data collected on flexor digitorum profondis could represent a valid measurement of objective difficulty, in our experimental tasks.

Objective difficulty, perceived difficulty, and ratings of difficulty.

The relationship between objective difficulty (defined as previously) and RSD was negatively accelerated. The best fitting had a logarithmic form (Figure 5):

24 paths, 15 subjects: RSDsc= (11.049)*logOD - (16.918) (r=.982)

27 paths, 6 subjects: RSDsc= (11.893)*logOD - (20.996) (r=.986)

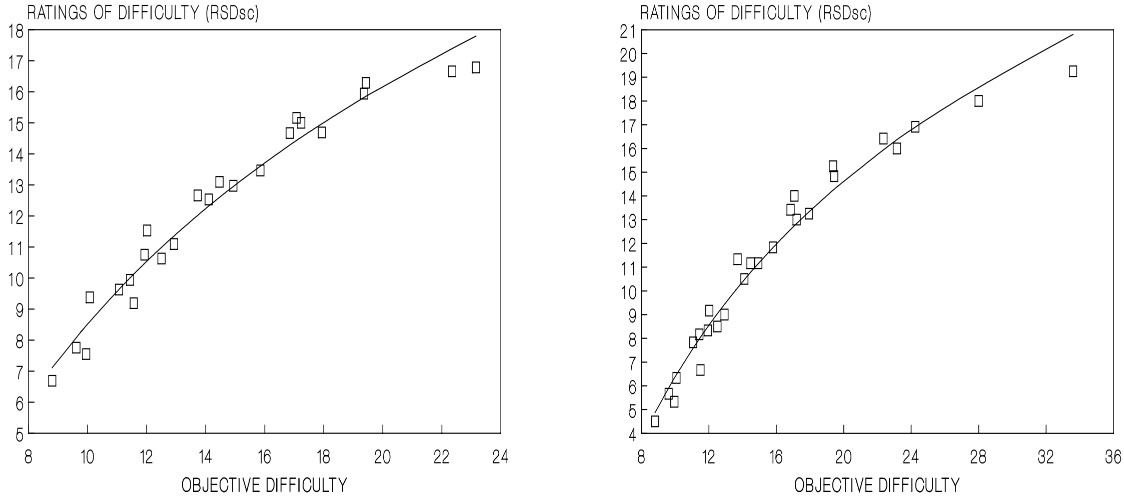

*Figure 5: Relation between objective difficulty and the rating scale of difficulty (scores). Left: 24 tasks, 15 subjetcs (RSDsc= (11.049)*logOD - (16.918), r=.982); right: 27 tasks, 6 subjects (RSDsc= (11.893)*logOD - (20.996), r=.986).*

The relationship between RSD and perceived difficulty was positively accelerated. The best fitting had an exponential form (Figure 6):

24 paths, 15 subjects: $PD= (3.932) * e^{(.140)RSDsc}$ (r=.986)

27 paths, 6 subjects: $PD= (6.010) * e^{(.114)RSDsc}$ (r=.978)

DISCUSSION.

Some results from this experiment could be dependent on the characteristics of our experimental device. It is the case for the relative contributions of exertion and accuracy in the estimation of difficulty, or for the importance of forearm muscles activity in the building up of perceived exertion. Another experimental device, giving more importance to technical paths, would probably give different results (Delignières, Famose, Thépaut-Mathieu et Fleurance, 1991). Presently, the discussion will focus on the results concerning RSD.

These results were generally consistent with our hypotheses. An important point was the improvement of rating accuracy, as difficulty increased: if the ratings proposed in the difficult paths appeared to be extremely reliable, our subjects seemed to have some problems to rate accurately easy tasks. Climbers seems to develop a specific expertise relative to the use of RSD, but this expertise is confined in the range of difficulties they use to practise. This confirms the idea that the judgement is elaborated by a comparison between the actual level of perceived difficulty, and a representation of the levels of perceived difficulty corresponding to each degree of the scale.

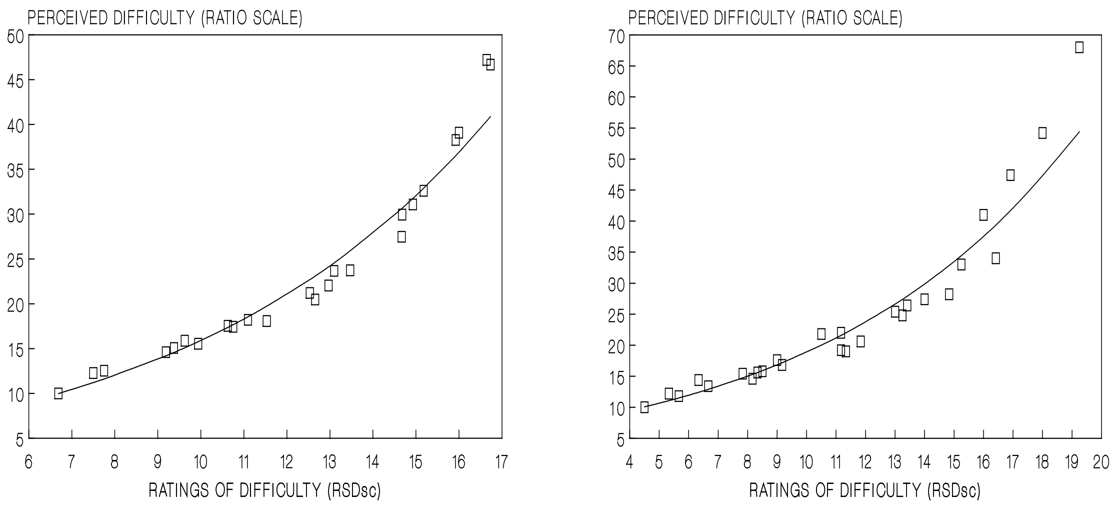

*Figure 6: Relationship between the rating scale of difficulty (scores) and perceived difficulty. Left: 24 tasks, 15 subjetcs (PD= (3.932) * $e^{(.140)RSDsc}$, r=.986); right: 27 tasks, 6 subjects (PD= (6.010) * $e^{(.114)RSDsc}$, r=.978).*

Another explanation could be proposed, related to the evolution of subject sensitivity, along the difficulty continuum: Blitz and Van Moorst (1978) showed using the signal detection method, that the differential threshold decreased when intensity of exertion increased. In other words, subjects perceived exertion variations better if they performed an intense exertion. This improvement of sensitivity could also explain, in our experiment, the fact that the discrimination between tasks seemed to be easier when the tasks were difficult.

This would incite climbers to caution when they rate paths and routes largely easy for them, in particular in situations of teaching or training. One may suppose that an individual recalibration of the rating scale, especially for the lower degrees, is necessary in the initial and continued formation of the trainers.

Our results indicated that RSD did not vary linearly with objective difficulty. The intervals of objective difficulty corresponding to each rate increased as difficulty increased. Climbers seemed to be more and more hesitant to use the upper rate. This is consistent with the hypothesis that the representation of one's upper limit introduce a context effect, as if the subject tried to preserve a reasonable margin between his actual difficulty rating and its maximal subjective limit. Then the judgment could represent a self-evaluation of the invested effort, and of the possibilities of surpassing. The compromise between the respect of this margin and the rise of the difficulty level introduce a ceiling effect in the relationship between objective difficulty and RSD. This property of the scale may explain in part the negatively accelerated evolution of the upper limit of the scale, during the last decade (1980: 7b, 1981, 7c, 1983, 8a, 1985: 8b, 1989: 8c).

Furthermore, the subjective interval between two successive ratings increased as difficulty increased. This was surprising, as it is generally assumed that the subjective intervals are equivalent all along the scale. Nevertheless, this could explain, in our experiment and in the common use of RSD, the multiplication of the intermediate ratings in the upper degrees: if for the lower degrees (2 and 3), our subjects used only two levels (e.g. 2 and 2+), they generally used three levels for the following degrees (5-, 5 et 5+), and six levels from the sixth degree (7a, 7a+, 7b, 7b+, 7c, 7c+). One may suppose that climbers will use even more discriminating systems for the future upper degrees.

Finally, the diverse relations we have described between RSD, objective and perceived difficulty are consistent with the results classically obtained in psychophysics: it has been shown that a category scale has a linear or logarithmic relation with the stimulus (Bonnet, 1986), and that for a given stimulus, a category scale and a ratio scale are related by an exponential function (Eisler, 1962; Ekman et Künnapas, 1962; Galanter et Messick, 1961; Stevens, 1968/69).

References

Blitz, P.S. & Van Moorst, A. (1978). Physical fatigue and the perception of differences in load: a signal detection approach. *Perceptual and Motor Skills, 46*, 779-790.

Bonnet, C. (1986). *Manuel pratique de psychophysique*. Paris: A. Colin.

Delignières, D. (1990). *La difficulté en escalade. Exigences objectives et perception des exigences des tâches motrices.* Mémoire pour le diplôme de l'INSEP.

Delignières, D. , Thépaut-Mathieu, C., Famose, J.P. & Fleurance, P. (1991). La perception des exigences dans les tâches de grimper. In J. Bilard & M. Durand (Eds), *Sport et Psychologie.* Dossiers EPS N°10.

Dornic, S., & Birbaumer, N. (1974). *Information overload and perceived difficulty in "neurotics".* Reports from the Institute of Applied Psychology, University of Stockholm, n°49.

Dupuy, C. (1987). *Analyse des stratégies visuelles et motrices d'une population de grimpeurs de haut-niveau.* Mémoire pour l'obtension du diplôme de l'INSEP.

Eisler, H. (1962). Empirical test of a model relating magnitude and category scales. *Scandinavian Journal of Psychology, 3*, 88-96.

Ekman, G. & Künnapas, T. (1962). Measurement of aesthetic value by "direct" and "indirect" methods. *Scandinavian Journal of Psychology, 3*, 33-39.

Galanter, E. & Messick, S. (1961). The relation between category and magnitude scales of loudness. *Psychological Bulletin, 68, 6*, 363-372.

Gonzales, M. (1988). Sur la caractérisation des processus cognitifs dans les modèles de jugement. In J.P. Caverni, C. Bastien, P. Mendelsohn & G. Tiberghien, *Psychologie cognitive, modèles et méthodes* (pp. 41-53). Grenoble: P.U.G.

Payne, J.W. (1982). Contingent decision behavior. *Psychological Bulletin, 92*, 382-402.

Stevens, S.S. (1957). On the psychophysical law. *Psychological Review, 64,* 153-181.

Stevens, S.S. (1968/69). Le quantitatif et la perception. *Bulletin de Psychologie, 276*, 696-704.

Tiberghien, G. (1984). *Initiation à la psychophysique*. Paris: PUF.