Published online: 10-1-2000

Performance Decrement Under Psychological Stress

Mitchell M. Berkun

Presidio of Monterey, California

Follow this and additional works at: <https://docs.lib.purdue.edu/jhpee>

Recommended Citation

Berkun, Mitchell M. (2000) "Performance Decrement Under Psychological Stress," *Journal of Human Performance in Extreme Environments*: Vol. 5 : Iss. 1 , Article 10.

DOI: 10.7771/2327-2937.1011

Available at: <https://docs.lib.purdue.edu/jhpee/vol5/iss1/10>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

This is an Open Access journal. This means that it uses a funding model that does not charge readers or their institutions for access. Readers may freely read, download, copy, distribute, print, search, or link to the full texts of articles. This journal is covered under the [CC BY-NC-ND license](#).

PERFORMANCE DECREMENT UNDER PSYCHOLOGICAL STRESS

MITCHELL M. BERKUN, *HUMAN RESOURCES RESEARCH OFFICE,
PRESIDIO OF MONTEREY, CALIFORNIA.*

Army trainees unaware that they were serving in an experiment were, under controlled conditions, led to believe that either (1) an aircraft in which they were passengers was about to make an emergency crash landing, (2) their outpost was now an artillery impact area, or (3) they had caused serious injury to a buddy by a mistake in wiring up explosive charges. They were required to demonstrate knowledge of crash landing procedures or repair a malfunctioning telephone to summon aid. Average performance under threat was significantly poorer than performance of the same act by other Ss motivated without threat. Soldiers with more military experience than trainees react differently both to the threat condition (they do better) and to the comparable non-threat condition (poorer).

Introduction

Relatively little is known about the average extent of performance decrement attributable to fear or psychological stress. For example, under combat conditions only the extreme or critical examples of good or poor performance ordinarily come to the attention of observers. When research is attempted in combat or in naturally-occurring disasters, any stress severe enough to equal combat in its effects can conceivably affect observers as much as the subjects, with some consequent impairment of the usefulness of their reports. Also, quantitative objective measures of performance effectiveness are not always possible in natural settings.

Laboratory experiments, for the most part, have either failed to tap this really severe source of behavior disturbance or have not measured effectiveness of performance of an act that is relevant to the stress. The present report offers a partial solution to the prob-

lem of quantification of response decrement under psychological stress.

The term "psychological stress" is meant to imply the presence of an apparent threat to the survival of the subject or of someone for whom he is responsible. While stimulus overload, heat, noise and vibration, difficult game decisions, and fine sensory discriminations obviously produce a condition which is frequently and reliably labelled "stressed," there is a basic motivation, drive, or attitude of *fear* not ordinarily manipulated in human factors research. It is argued that the unknown contribution of this variable jeopardizes the interpretation of data from laboratory studies of "stress." Subjects who are aware that they are subjects, that they are being tested, and that they are, quite naturally, in the protection of the responsible experimenters may be quite different individuals from the same subjects responding to unexpected threats to life and limb. For example, during preliminary experimental attempts at HumRRO to produce a combat-like state of fear in normal limit subjects, a defensive phenomenon was seen to develop, even when there may have been, objectively viewed, a real threat to our subjects. In one case, there was a chance of falling from a high, unstable bridge; here the subjects denied to themselves and to us that they would be deliberately exposed to real danger. They usually realized, quite quickly, that they were in an experiment. They assumed that they were safe and that they were sometimes expected to act scared. Their perception of these experimental situations was that the likelihood of danger was very small, that there was no reason to actually be scared, but that the experimenter frequently expected them to show upset or fright.

We have termed this phenomenon "cognitive defense" because its essential element is the complex act of cognition, thinking, assessing probabilities, and using the most likely event as a

basis for action. This characteristically human cognitive function, then, becomes the principal obstacle to experimental study of the response to stress. This denial of threat during an experiment is obviously justified, as it is not ordinarily justified when it occurs during natural disasters.

We need to induce in the subject a state which has very much in common with the state that is induced by the extreme environments that we maybe training people to cope with. To achieve this desired type of stress, it is necessary to overcome the "cognitive defense," to recreate in subjects those elements of naturally-occurring disasters that we do have the fear-producing effect. Then we can know the contribution which this fear component makes to effectiveness and persistence of performance in stress. The cognitive defense can be overcome by using a "cognitive stress stimulus," a stimulus that must be thought about by the subject in order to have the desired effect. While perpetual supports (or stage props, which is what they really are) are required, the essential element is the *cognition* whereby the subject figures out that he is in trouble. His cognition, then, is a matrix in which embedded the heat, the cold, the stimulus overload, the abnormal gravity and whatever else modern technological mastery is going to expose him to. Without this cognitive basis, the perpetual supports are denied out of existence by experimental subjects. The function of the cognitive stimulus is simply to preserve the natural fear-producing effect of the stage props in the face of such denial.

The following four requirements are stated for an experimental research program on the determinants of the normal or modal person's ability to cope with stress. First, it is necessary to measure the performance acts relevant to the stressful environment. Second, there must be an objective measurement of this performance level. Third, it is necessary to allow for a possible difference in the effect of serious threats to life as compared with the simple effect of information overload, achievement failure, or other laboratory stresses. Fourth, the effect of test-taking set or an experimenter-oriented motivation on the subject's part must be controlled (Orne, 1962).

An estimate will be presented here of the extent of behavior degradation of normal persons in experimental situations based on the above reasoning. The experimental situations from which these data derive will be shown to meet certain rigorous criteria of stressfulness, and the possibility will be raised of their being useful and practical for further research along lines important to the reader. These situations were created and tested by the staff of a continuing project known within HumRRO as Task FIGHTER, whose goal is the understanding of motivational factors contributing to effectiveness under fire in combat. The principle characteristic of these situations is the presentation to the subject, under controlled conditions in which the subject is actually safe, of an apparently real emergency in which human life is at stake. The effects of other stresses such as fatigue, heat, noise and hunger, are deliberately minimized in these studies. More discussion of certain aspects of these situations is given by Berkun, *et al.* (1962).

The First Situation: "Ditching"

Subjects, all young men early in their basic combat training, were told they had been recruited to participate in a psychological study of the effect of altitude on thinking. They were taken up, 10 at a time, in a DC-3 aircraft fitted as a flying classroom, and at 5000 ft were given a simple arithmetic reasoning test. All four

panelloads had the same mean scores on this test, so that subsequent test differences between experimental and control subjects could, in all likelihood, be attributed to the difference in the way they were treated.

Before boarding the aircraft, all subjects were put through what appeared to be SOP for military flights. They spent 10 min reading a manual on ditching procedure, they donned Mae Wests and parachutes, and they met the "steward" who was introduced as the aircraft commander's representative in charge of the cabin and its passengers. The subjects and the apparent experimenter accepted the steward's authority. It was not revealed to the subjects that he was also, in fact, a psychologist and one of the experimenters.

The two control flights changed altitude uneventfully after the first test, and the "steward" briefly took over from the "experimenter" to have all passengers fill out two ostensibly routine questionnaires. The first, Emergency Data Form, purports to be a reckoning of the subject's personal possessions, and is a deliberate example of really bad human engineering. It is scored for accuracy and consistency in following complex instructions. The second is an achievement test measuring retention of the ditching procedures information read before the flight. Subjects were not aware that these were tests, and accepted them as part of the Army's way of doing things. The flight terminated shortly after these tasks were accomplished.

Two flights were made with passengers who had been randomly assigned to the experimental group. On these flights the change of altitude following the reasoning test was accompanied by rolling and yawing of the aircraft, sputtering of an engine, and the stopping of one propeller, all under the deliberate control of the pilot. Subjects "accidentally" overheard a pilot-to-tower conversation involving an emergency landing, and had every reason to believe this was a genuine conversation. At this point the "steward" assumed control of the passengers from the "experimenter" and announced a state of emergency. Both the subjects and the experimenter were required to fill out the two forms just described, with the understanding that this was required procedure under the emergency conditions.

Of the 20 subjects exposed to this stress of a threatened ditching or crash landing, five reported that they had detected the hoax, and their results were deleted. For example, two were experienced pilots, and knew the capabilities of the DC-3. Another noticed, while he was multiplying his age by 12 to get the age in months required for the Data Form, that a previous occupant of his desk had used the airsickness bag as scratch paper to do the same computation. From this he figured out that something systematic was going on, thus cognitively concluding that the plane was not in danger. These are typical of the problems run into in this line of research.

The remaining 15 experimental subjects scored a mean of only 35 out of a possible 50 points for correctly completing the Emergency Data Form, a mean significantly lower than the 39 point mean of the 20 control subjects. This can be stated as a 10% decrement in cognitive verbal behavior under stress as compared with unstressed subjects.

Retention behavior suffered even more. The control subjects correctly recalled an average of 7.6 of the 12 answers; only 4.9 were recalled correctly by the stressed subjects. Again, this difference is statistically significant, the decrement being about 28%. Observe that the greater decrement seems to occur on this task

which is clearly more related to survival than is the "Data Form" task.

The Second Situation: "Artillery"

Subjects, basic trainees again, were transported to an isolated semi-wilderness area, where they were told they were to participate in a large-scale military exercise. The purpose of the experiment, they were told, was to test the concept of mobilizing raw troops who have not completed conventional basic training for combat. Large numbers of troops were to be participating and artillery, chemical and nuclear aspects of war were also to be involved in the maneuvers.

Each subject was taken individually along a devious, winding route to an isolated outpost where he was stationed with a two-way radio. This was supposed to be an experimental model of an advanced transceiver being tested in the field for the first time.

After briefing him on using this radio to report any aircraft flying overhead, his escort departed. Subject was then assigned the call letters "Bravo Five" by the Command Post (CP), and could hear Command Post's messages for other outposts seemingly just like his own. After 25 min of normal procedure, partly remarks of the CP radio man, and partly messages having to do with routine administration of this military exercise, the tone of voice of CP changed. Reports were coming in of artillery rounds bursting outside the designated impact areas, and these reports were passed on to "Bravo Five" as well as to the fictitious other outposts. At about this time a 5-lb charge of TNT buried some 150 yd from the subject was detonated, simulating the burst of an incoming artillery round. At 45 min after first being stationed at his radio, the subject heard this taped message: "Bravo Five, this is Command Post. I've received no transmissions from you in quite some time. I suspect your transmitter's gone out - convert to emergency transmission." Twenty more minutes, and he heard that the experiment was terminated because of the danger. It was urgent that he repair his transmitter in order that he be located, because his position had been lost by Command Post and there was no other way (apparently) for him to be rescued from the impact area. During this period six more simulated shells burst, increasingly closer to him.

In fact, the subject was in constant communication with CP but he did not know this. When he began to follow instructions posted on his "radio" for conversion to "emergency transmission" he unknowingly took a performance test set up for him. He believed he was doing something to bring about his rescue from a very dangerous situation. The actual competence he demonstrated could be measured objectively. A number of scores, representing the speeds with which various subtasks on this repair job were accurately completed, are consolidated into a single measure. The range of this consolidated performance score is from 0 to 10 points; it includes the heaviest weight for completion at all, with relative lesser weights for relative speeds of completion.

Subjects in the control group (17 individuals) participated in the same exercise, but for them nothing serious went wrong. They were also addressed as "Bravo Five," and the knowledge of their position also was lost by Command Post. There developed a small problem of delivering future water and fuel supplies to "Bravo Five," but as he was given this well-fed this was not at all critical. However, his "transmitter" failed, and he was requested to convert it to the stand-by transmitter circuit. Thus, he also took the test unaware that he was being tested; but his motivation, unlike the

experimental subject's, definitely was not a matter of life and death.

There are no observed differences in intelligence, aptitude, or education between the experimental and the control subjects to contaminate any differences found in performance.

The control subjects earned a mean score of 6.05 points out of the theoretical maximum of 10 points for speed and accuracy of the repair task. This is our standard of comparison.

The 24 subjects with an apparently real and certainly vital reason for fixing the "radio" to be taken out of the artillery impact area earned an average of only 4.05 points. This decrement which amounts to a 33% drop from standard performance, is statistically very significant. Contributing to this mean decrement is the defection of 8 of these 24 soldiers. These individuals abandoned their posts at various times in defiance of specific orders from CP to remain and repair their radio sets. Since they did not attempt to complete the circuit conversion, their scores were lower and brought the group mean down. There were no control defections. Removing these 8 zero scores still leaves the experimental mean below the control mean, but only at a borderline level of significance.

The Third Situation: "Demolitions"

In the two previously-mentioned stress situations, the subject himself was threatened with death. Here he finds himself the unwilling and unwitting instrument for serious injury or death to another.

A work detail was brought to a remote area of Fort Ord, California, and its members isolated from one another. One at a time, each man was taken by an experimenter to an isolated bunker and was told to screw colored wires onto matching terminals in order to install a remote control switchbox for explosives being detonated in a nearby canyon. He was left alone, he thought, when all other personnel involved in this construction project had gone down into the canyon out of his sight. He had been shown how to crank, talk, and listen on an Army field telephone switchboard in the shed where he was left, and was cautioned to listen for instructions on a conventional intercom or "squawk box" also in the shed. In the next 3-5 min the subject completed the wiring and threw a test switch as instructed. This set off a 5-lb charge of TNT down in the canyon. The explosion rocked the bunker, sometimes shaking things off shelves. Then a voice came over the intercom. It was a tape recording, but the subject did not know this. He heard:

"Listen, if you can hear me, we had an explosion down here and I think someone was hurt. I want you to stay right there.

Listen, if you can hear me, wait right there and don't touch anything. Listen, are you sure you did that wiring right?"

Three minutes later the subject heard a taped request that he telephone the post hospital for medical assistance. But the phone will not work properly and he is told to follow instructions posted on it for restoring it to service. Several other recorded messages followed over the next 45 min, concerned with the subject's progress on the telephone, the difficulty of keeping the injured man alive, a reference to the fact that the Military Police will want to question the subject, and a final message that someone was on the way up to question him.

The "telephone" was essentially the same as the "radio" which subjects in the "Artillery" problem were required to fix, and the same consolidated performance measure was recorded. This score

represents competence during the problem, speedily and accurately doing a job required by circumstance.

Another measure, not at all related to the problem that actually (as he saw it) faced the subject, was taken after "help" arrived at the telephone. The subject was asked to repeat in reversed order a series of random digits recited to him. As in the previous study, the control subjects had to "repair" the same equipment for a less critical but still valid reason: they were to call Post Quartermaster to arrange for future delivery of rations to the area. Again, the two groups of subjects are equally matched on intelligence, aptitude, and education.

Out of 10 possible points for telephone repair, the control group earned a mean of 0.10, or very close to the control mean in the previous experiment. The experimental group earned a mean of 5.01, significantly poorer than the controls. The response decrement attributable to the stress is 18% of the control value.

On the Backward Digit Memory Span tested after the stress ended, the controls, on the average, were able to correctly recall a series of five digits backwards; the stressed subjects could not recall a five-digit series, but could recall a four-digit series backwards. The difference in the two means is statistically significant. Subjects appeared to accept this test as part of the situation, not noticing any logical difficulties in accounting for the task.

An 18% competence decrement in performance under stress thus appears both in a completely irrelevant immediate-memory test administered after the stress ended as well as in a relevant task during the stress. This decrement is about as large in this situation where the stress involves only the survival of another person as it was in the other experiments where the subject's own survival was at stake.

Do these Situations Actually Produce Stress?

Examination of the published psychological literature on stress, discussion with military consultants, and the study of subjects in preliminary experiments resulted in the statement of three operational criteria for experimental stressful situations. These criteria were set up before any of the experiments reported herein were begun.

A situation or problem intended to be perceived by a subject as threatening or stressful must affect the subject in the following three ways:

Performance

It must result in behavior which can be shown statistically to differ from behavior of persons attempting the same task but without having to cope with a comparable stress. Fig. 1 summarizes this outcome; all contrasts are statistically significant.

Subjective Feelings

Subjects must report feeling distressed in the situation. To quantify this response each subject is offered a list (Kerle and Bialek, 1958) of words that have previously been scaled by a judging procedure, and is asked to select the one word that best describes his feelings at the time. His score is the scale value of the word chosen. The "indifferent" point is 48; higher scores indicate feelings of greater distress. Extensive experience (Berkun, *et al.*, 1958) with this Subjective Stress Scale (SSS) has demonstrated it to be consistent, sensitive, reliable and valid. Fig. 2 gives the scale and its results. All three situations meet this criterion also.

Figure 1. Performance levels in the three stress experiments as percentages of maximum possible score in each.

Figure 2. Mean quantified subjective verbal self-report (SSS) in each stress experiment.

Physiology

There must be some indication either in blood or the urine of a disruption in normal body processes. It has been shown that physical stressors cause an excessive activity of the adrenal cortex. This in turn affects the presence of a particular class of substances in the urine the 17-hydroxy-cortico-steroids, and affects the count of a particular kind of white cell (eosinophil cells) circulating in the blood. The subject exposed to situations intended to elicit stress must show a difference in one or the other of these measures, as compared to control subjects on the same diet tested at the same time of day. In all three reported experimental groups either the eosinophil count was lowered or the urinary steroid level was raised, or both. Because these experimental levels differed significantly from control levels, according to statistical tests, it is concluded that the experimental situations did indeed disrupt normal body chemistry.

Two other situations, not previously discussed, were evaluated in the same general set-up as was the "Artillery" problem. They threatened the subject with death from a forest fire or front accidental nuclear radiation or fallout. Both of these failed to meet all three stress criteria, presumably because the artificial smoke and the dummy dosimeters failed to adequately support the information received by the subject from his Command Post. However, the three situations described herein met all three criteria.

The Effect of Longer Military Experience Prior to becoming a Subject in these Experiments

As part of the "Artillery" experiment, an additional group of subjects was obtained who, in contrast to all others in this entire series of experiments, had been serving in combat-ready units and in a special experimental regiment participating in exercises testing new concepts of warfare. The experienced troops were older than the recruits, but had had less formal education and had lower scores on several aptitude and intelligence tests from the Army Classification Battery.

These experienced subjects responded somewhat differently from the relatively inexperienced troops, both in their competence at fixing the radio and in their subjective self-evaluation. This difference from inexperienced troops is complex, as it is itself a function of whether they were assigned to the experimental stress or to the control condition. Fig. 3 presents the picture for performance scores. The line in Fig. 3 labeled "raw recruits" simply repeats the data reported earlier showing response decrement under stress. In contrast, the reader will observe that the experienced troops do better under stress, poorer in the control condition. This crossing-over shown in the diagram meets the test of statistical significance.

Likewise in Fig. 4, there is a statistically reliable crossing over of the two lines. A higher mean distress score (paralleling the lower performance score) in the stress condition characterizes the raw recruits. But, again, the experienced troops who were seen in Fig. 3 to perform poorer in the control condition than under stress show a corresponding increase in the intensity of their feelings of distress.

It is reasonable to attribute the performance decrement and negative feelings of the experienced troops in the control conditions to a form of boredom and disappointment when required to do what they think of as only "busywork." There are indications from this and earlier unpublished studies that certain bland conditions intended to be nonstressful become, in fact, mildly stressful

Figure 3. Mean performance scores in the "Artillery" situation.

Figure 4. Mean subjective self-report in the "Artillery" situation.

for such subjects. When, however, the situation becomes personally important, then their maturity and experience are brought into play, and they cope with their environment somewhat better than do their untrained counterparts.

DIFFERENCES BETWEEN MORE COMPETENT AND LESS COMPETENT PERSONS

All subjects, experimentals and controls, were dichotomized at their respective group median on the performance test. The Classification Inventory, a sub-test of the Army Classification Battery, was administered to the subjects in the "Artillery" and "Demolitions" situations. This test is an autobiographical personality interest attitude and activities questionnaire. The better performers, that is, those whose speed-accuracy score was above median, earned a significantly higher average score than did the below-median performers. The same result was found with the Taylor Manifest Anxiety Scale: the poor performers were more anxious persons. In general these have achieved less in their earlier lives, had less general aptitude, less self-confidence and lower intelligence than their better-performing counterparts.

DISCUSSION

That all the situations tested, including the two not meeting all our *a priori* criteria, uniformly produced a mean decrement in the performance of naive subjects was, frankly, a surprise to the experimenters. The *a priori* statement of the performance criterion had allowed for any change in the distribution of scores, and did not predicate a mean difference in either direction between the experimental and the control group.

The findings with the experienced group of subjects are particularly significant for human factors research, since they present the possibility of a greater *ennui* factor for sophisticated than for naive subjects. To state this in another way, the experienced subjects more readily discriminate some "play" situations from the "real thing." They can, of course, "play" more intensely, as documented by common observations in aircraft emergency training simulators, but it is conceivable that the subjective feeling of stress and the decrement in response efficiency observed in simulators may not necessarily generalize to actual emergencies aloft. Since the exact nature of the relationship of this "play" effect to experience is not known, there is some risk in applying this consideration to simulator research at the present time. Obviously, however, it bears exploration.

Some other, less objective observations on behavior in these life-threat situations are worth mentioning here. The subjects underwent a severe restriction of the perceptual field. Relevant stimuli were not noticed, and inadvertent cues which the experimenters feared would compromise the deception failed to "register" with the subjects. Whole graphs from the instruction manual supplied with the "telephone" were skipped, and subjects subsequently denied that the information therein was available at all. Some subjects either did not see the screw driver provided, or actually recalled seeing it yet drawing no connection between it and the screw tightening job they were trying to do on the switchboard.

Some subjects reported becoming engrossed in the repair job to the exclusion of other considerations, so that they resisted the interference of the experimenter coming to their rescue and so that they temporarily forgot the purpose of their hard work on the

switchboard. The drive-reducing aspect of burying oneself in this kind of work is not to be ignored in systems designing.

CONCLUSION

To predict from experimentation the ability of men to cope with real stresses requires, first, a validation of the experimental situation as a substitute criterion for uncontrollable reality. It is argued that the subject must cognitively perceive the situation as stressful, so that he may react realistically and not "as if." Simulation of a stressful environment, then, must avoid cues which invite the subject to deliberately assume a role or which provide him with more psychological support or sustenance than he will receive in the reality to which the findings must generalize. Furthermore, the task he is to perform must be meaningful in the stress-producing context. Stressors which fulfill these requirements ought to produce (1) a measurable disturbance of performance, (2) a report of awareness of a feeling of discomfort, fear, threat, or unpleasantness, and (3) a measurable perturbation of physiological (homeostatic) processes. Some experimental situations involving apparent accidents perceived as real by the subjects have proven feasible and successful in producing a group effect on the above three dimensions. Before the situations were tested, there was no prediction of the direction of disturbance of the task performance or of the physiological measures. In fact, all situations subsequently were characterized by a degradation of speed and accuracy. The data also indicate a difference between naive subjects and those more familiar with the context in which the "accident" occurred.

REFERENCES

- Berkun, M. M., *et al.*, Experimental studies of psychological stress in man, *Psychol. Monogr.*, 1962, whole No. 534, 39 pp.
- Berkun, M. M., Timeras, P. S., & Pace, N. Psychological and physiological responses in observers of an atomic test shot. *Psychol. Rep.*, 1958, 4, 679-682.
- Kerle, R. H., & Bialek, H. M., *The construction, validation and application of a subjective stress scale*. Staff Memorandum, Feb. 1958, U.S.A.L.H.R.U., Presidio of Monterey, California.
- Orne, M. T. On the social psychology of the psychological experiment. *Amer. Psychologist*, 1962, 17, 776-783.