EVALUATION OF THE ALLEGED REDUCED NARCOSIS OF NITROX MIXTURES

Erez Heilweil

Thesis submitted in partial fulfillment of the requirements for
the Master Degree

To the
University of Haifa
Faculty of Humanities
Department of Maritime Civilizations

January 2007

EVALUATION OF THE ALLEGED REDUCED NARCOSIS OF NITROX MIXTURES

By

Erez Heilweil

Supervised by: Dani Kerem, Ph.D.

and Limor Sheffer, Ph.D.

Thesis submitted in partial fulfillment of the requirements for
the Master Degree

To the
University of Haifa
Faculty of Humanities
Department of Maritime Civilizations

January 2007

Approved by: ______________________________ Date: ____________________

(Supervisor)

*Approved by: ______________________________ Date: ____________________

(Supervisor)

Approved by: ______________________________ Date: ____________________

(Chairperson of M.A. Committee)

THANK YOU

First of all I would like to thank my wife Ayelet, without whose support and sacrifice this thesis would not have come out the way it did.

I would also like to thank my two supervisors, Dr. Dani Kerem who provided me with tools to wirte this thesis and Dr. Limor Sheffer who took me under her wings and spent long hours sharing with me her wide knowledge and providing me with psychological tools.

I also want to send my appreciation to the Israeli National Foundation for Recreational Diving Develepment, the Jacob Recanati Foundation and the Maurice and Ireine Hatter Foundation who financed this research.

I am very grateful to the entire staff of the Recanati Institute for Maritime Studies who have fulfilled all my needs without any hesitation, and to the wonderful research diving crew, Dor Edellist, Omri Ben Elihu and Nir Hadar, led by Stephen Breitstein and Amir Yurman.

A special thanks to the Maritime College in Mikhmoret and "Mevoot Yam" Marine School for providing inshore and offshore facilities for this thesis.

Finally, I would like to thank all of the participants in this experiment who dove regardless of uncomfortable conditions such as the water temperature.

Thank you all…

TABLE OF CONTENTS

Evaluation of the alleged reduced narcosis of nitrox mixtures
Erez Heilweil

ABSTRACT

Background

Research shows that success or failure in performance is influenced not only by one's abilities and limitations but also by one's judgment. Nowadays it is widely acknowledged that one tends to act according to one's own judgment whether or not it is the right thing to do. This fact may turn against us, all the more so when performing underwater tasks under the influence of nitrogen narcosis.

Nitrogen narcosis is a phenomenon similar to alcohol intoxication, caused by high partial pressures of nitrogen that cause its excessive dissolution in the membranes of the neurons in the central nervous system. The symptoms cover a wide range of severity, starting from mild impairment of performance on unpracticed tasks (at 30 meters of sea water (msw)), up to hallucinations and general anesthesia (at 90 msw and deeper). The most hazardous outcome of nitrogen narcosis is not that of extreme symptomatology, a rare event in cases of uncontrolled descent to great depths, but rather the stages of euphoria, overconfidence and loss of judgment which occur at intermediate depths, and which may cause the diver to lose alertness, take extra risks and start a chain of events culminating in a serious diving accident or loss of his/her life and/or that of his/her diving buddy. In addition, even subtle decrements in mental performance occurring at moderate depths, especially when unrecognized by the diver, may compromise intricate underwater missions, both military and scientific.

Nitrox mixtures are obtained by adding oxygen to air, and are known as Enriched Air Nitrox (EANx) or nitrox_x, the x denoting the percentage of oxygen in the mixture, i.e. $\text{EAN}_{36,}$ nitrox_{36}. The main advantage of nitrox mixtures is a reduced tissue nitrogen load relative to air at the end of a given dive profile (an advantage that merited the name SafeAir for EAN). This allows increased bottom time and reduces the risk of decompression sickness. Another benefit centers on the claim that nitrox mixtures are less narcotic than air at any given depth. The assumption behind this claim, which so far has not received rigorous experimental support, is that the high partial pressure of

nitrogen is the sole cause of the effect. However, a high partial pressure of oxygen, as found in nitrox mixture, might totally or partially compensate for the loss of nitrogen, since oxygen is twice as lipid-soluble as nitrogen.

This study attempts to quantify nitrogen narcosis during open sea sport diving at the moderate depth of 30 msw, as related to the type of breathing mixture, nitrox or air. In addition, it examines the influence of the breathing mixture on the diver's ability to correctly judge (both predict and assess) the decrement in his performance caused by nitrogen narcosis. The experimental approach was based on a combination of cognitive tests and metacognitive measures, not previously applied in the undersea environment. Sensitive cognitive tests served to quantify performance decrement, and metacognitive tools were used to quantify the quality of self-judgment and the matching of self-confidence with actual performance. The combination of these allowed the overall mental ability of the diver as a function of breathing mixture to be assessed. Performance and judgment were examined at a depth of 30 msw for both air (30A) and nitrox$_{36}$ (30N) against a control dive at 2 msw with air (2A) – the nature of the mixtures being irrelevant at this depth.

The main objective was to verify or refute the alleged reduced narcosis of nitrox$_{36}$ over air. Secondary objectives were determining whether narcosis affected the encoding or the recall phase of new information, or both, quantifying the often described over-confidence induced by nitrogen narcosis, and determining whether it is affected by the breathing mixture, and finally, examining the consistency over time of nitrogen narcosis effects on the individual diver.

This research was approved by the Ethics Committee of the University of Haifa (Helsinki Committee).

Methodology

The experiments were carried out at the Israeli Maritime College in Mikhmoret, at two flat-bottomed seabed locations chosen to provide the desired depths. The marina enabled the 2 msw (shallow control) dives, and a buoy anchored 2.4 kilometers off shore marked the site of the 30 msw dives (deep test dives).

Equipment: The equipment consisted of personal diving gear, nitrox-compatible single hose scuba regulators, nitrox-marked cylinders filled with air or nitrox$_{36}$ (the contents of which were only known to the diving officers), safety equipment including a nitrox_{50} cylinder attached to a safety-stop platform at five meters' depth, a pure oxygen cylinder and a diving flag on board the boat, a laptop, polyester A4 paper (DFDM100), rough Teflon writing slates, pencils, carrying crates, a timekeeping slate and an underwater horn. Transportation to the offshore dive site was by speedboat, which was also used as a gearing up platform and a safety boat.

Participants and experimenters: The participants in this experiment were 35 divers, 23 male and 12 female, between the ages of 20 and 30, with a diving history range of 13 to 44 dives. All were Hebrew-speaking undergraduate students at the Israeli Maritime College in Mikhmoret, who were paid per dive for participating in the experiment. All signed a consent form and a medical declaration, and presented a diving certificate and updated insurance. Due the tight schedule and study constraints, not all 35 participants completed all scheduled dives (as detailed below).

This study was managed by a team of experimenters headed by the experiment manager. Task allocations were as follows. The Experiment Manager was in charge of constructing the questionnaires, monitoring the diving schedule (depth, questionnaire). The Diving Supervisor was in charge of filling the cylinders, checking the dive mixtures, providing diving equipment and keeping track of safety issues. The Timekeeper was in charge of timing the phases of the experiment, as well as being the safety diver. The Tester was an advanced diver (dive master or scuba instructor), whose role was to act as a dive buddy to the participant, carry, hand out and collect the questionnaires; The Skipper was in charge of the boat and the safety swimmer.

Experiment design: The experiments were double-blind (neither the participants nor the testers knew the nature of the deep gas mixture): therefore, no-decompression limits were calculated for air, causing bottom time not to exceed 20 minutes. The experiment was of a self-control type, each diver making two 30 msw dives, one for each mixture, and a 2 msw dive which controlled the effect of submergence *per se*. Dive order assignments were balanced by a Latin square.

Tests: In each of the three dives, three kinds of memory tests were given to the participants (a different version in each dive). Each test consisted of two phases: encoding phase and recall phase. Twenty-three divers completed the 'Faces Test' (Experiment 1), designed to assess depth effects on retrieval of visual episodic memory. The encoding phase was performed on land, where divers were required to memorize 30 faces in a fixed time, and for each one, note their level of assurance in their ability to recall it underwater, when presented together with three unknown faces. The assessment was a value between 25% (guessing) to 100% (absolute certainty). This assessment is a metacognitive measure, known as judgment of learning (JOL). After reaching the sea bottom, the recall phase began, in which participants were asked to identify the memorized face among three unfamiliar faces, and for each one, note their degree of confidence (25%-100%) in the correctness of their answer. This assessment too is a metacognitive measure, known as confidence.

The 'Word-pair Test' (Experiment 2) was performed by 26 divers. Its purpose was to assess both encoding and recall of textual episodic memory underwater. In the encoding phase, divers were exposed for a fixed time to a list of 50 word pairs which they were instructed to memorize for free recall of the second word (rather than choosing it from a list of 4) when presented with the first. JOL was given, on a scale of 0%-100%. The recall phase was not followed by a confidence judgment, as the time delay from the JOL was deemed too short.

The 'Picture Test' (Experiment 3) was intended to test the effect of depth/narcosis on the encoding phase of visual episodic memory by exposing the diver to a picture underwater (plus a single overall JOL score), and then asking him/her to recall details from it when back on shore, through multiple-choice questions with confidence option (on a scale of 25-100%).. The results were later revealed to be biased due to technical problems, and therefore this experiment was cancelled.

The JOL and the confidence are metacognitive measures from which we can derive the calibration and the resolution of the participants. Calibration is a crude estimate of the match between judgment (either JOL or confidence) and actual performance of the participant, calculated by subtracting the correct score (%) from the JOL or confidence (a negative score is defined as under-confidence, a positive score means over-confidence and zero means perfect judgment/evaluation). Resolution measures the ability of the participant to correctly recognize the questions on which he/she is more likely to err, or on which he/she actually erred. It does so by two indices: discrimination index and gamma correlation. The former, with a range of -100 to +100, is the difference between the mean (JOL or confidence) score given to questions with correct answers, and that of questions with wrong answers. Gamma correlation, ranging between -1 and +1, is a finer test, taking account of the scores given to each individual question.

Experimental protocol: The experimental procedure began with the encoding phase of the Faces Test (Experiment 1), followed by equipment organization, gearing up and diving to the sea floor (2 msw or 30 msw). This process took between 40 to 60 minutes. The recall process of Experiment 1 took 2 minutes, followed by the encoding of Experiment 2, and after 30 seconds, the recall phase of Experiment 2. Experiment 3 started immediately after Experiment 2 by a 30 seconds encoding of the picture. Half an hour later, after reaching the shore, the recall phase of Experiment 3 was executed. This procedure was performed three times per participant, once for each of the conditions (2A, 30A and 30N), spaced by a period of 1 to 42 days

Main results

Experiment 1 showed a striking resemblance between all three dives. Depth did not affect recall performance or confidence, and obviously, there was no effect of the breathing mixture. Group mean JOL and confidence gave near-perfect calibration, demonstrating neither under- nor over-confidence. In contrast, Experiment 2 demonstrated a clear decrease in performance at depth. This decrease was not accompanied by a lower JOL, resulting in a positive mean group calibration value, indicating over-confidence. There were no significant differences between nitrox and air.

There was a significant positive correlation between individual cognitive performances, as well as in the metacognitive measures between experiments, and for the two deep dives of Experiment 2

Discussion

Depth effects: The results prove that even at the shallow depth of 30 msw, the Word-pair Test was sensitive enough to reveal a deterioration of performance not backed by a reduction of JOL, meaning the participants as a group were not able to foretell their impaired ability at depth. The resulting positive calibration value, indicative of overconfidence, has been mentioned in the literature, but never previously quantified. One explanation for the apparent lack of depth effect in Experiment 1 is that, being recognition rather than a free-recall type, it was too easy. Another possibility is that it is the encoding rather than the recall phase that is negatively affected by narcosis. If this is the case, this finding has never been described before.

Dive mixture: The similarity in results between nitrox and air concerning the nitrogen narcosis effect may be interpreted in several ways: 1. The two mixtures have an identical narcotic potency at this depth; 2. the 30 msw depth is too shallow and/or the sample size is too small to bring out an existing effect; 3. the impaired performance is not due to narcosis, but to other depth factors such as cold, visibility or anxiety.

Screening tool: One potential practical outcome of this research would be to use the tests as a basis for screening candidates for professional deep diving. A necessary condition for this, that seems to exist according to the results, is a strong correlation between personal performance and judgment during different dives. This finding strengthens the contention that both surface performance and judgment and their modulation by depth/narcosis are time-stable traits. Therefore, applying the tests and analyzing individual performance and self-appraisal may enable weeding out candidates possessing a high vulnerability to narcosis and a low awareness of this fact.

Summary

The primary aim of this research was to determine whether nitrox diving involves a lesser degree of narcosis at a given depth. Whether the apparent negative answer will hold for a larger sample and for a greater depth range will only be revealed by further research. Overall, this preliminary research was a pilot and probe for future studies using the cognitive-metacognitive combination approach. This was why it was constructed in a manner that could evaluate different experimental tests and conditions. The metacognitive measures, used for the first time under water, provide a new dimension to the study of the nitrogen narcosis effect

It is clear that Experiment 2 gave the better results and that with time constraints, should be the one used for the entire experiment, i.e., participants should perform the word-pair encoding and recall, including JOL and confidence, in all three experimental versions. However, since visual memory plays an important role in undersea tasks, the Picture Test should also receive a similar evaluation.

The results show a clear manifestation of over-confidence at depth, i.e., there is a deterioration of performance, which is not matched by a reduced JOL. There is also no discernable effect of the gas mixture, and it can be claimed that at 30 msw depth both mixtures have the same narcotic effect. The consistency in results over time indicates that Experiment 2 was reliable, and that there seems

to be an innate individual sensitivity to nitrogen narcosis, which, if verified by further studies, may help construct a screening test for professional deep divers.

Keywords

Cognition; individual differences, metacognition; enriched air nitrox; nitrogen narcosis; scuba diving; gas mixture; diver selection

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: BACKGROUND

1.1. <u>History</u>

The desire to go under water has probably always existed: to hunt for food and other riches of the sea, to uncover artifacts, to repair ships (or sink them!), and perhaps just to observe marine life. Until man harnessed technology to invent means of breathing under water, however, dives were short and shallow. Human diving can be divided into three stages (Bachrach, 1982):

1.1.1. Free or breath-hold diving

Man is known to have practiced breath-hold diving for millennia. Indirect evidence comes from undersea artifacts found on land (e.g. mother-of-pearl ornaments from 4500 BC), and depictions of divers in ancient drawings. In ancient Greece, breath-hold divers are known to have hunted for sponges and engaged in military exploits. Of the latter, the story of Scyllis (about 500 BC), as told by the 5th century BC historian Herodotus is perhaps the most famous. During a naval campaign, the Greek Scyllis was taken aboard ship as prisoner by the Persian King Xerxes I. When Scyllis learned that Xerxes was to attack a Greek flotilla, he seized a knife and jumped overboard. The Persians could not find him in the water and presumed he had drowned. Scyllis surfaced at night and cut the moorings of all the ships in Xerxes' fleet, using a hollow reed as a snorkel to remain unobserved. Then he swam 9 miles (15 kilometers) to rejoin the Greeks off Cape Artemisia.

1.1.2. Diving bells and helmets

Alexander the Great's descent in the primitive diving bell *Colimpha* (about 330 BC) is one of the earliest reports of a diving bell. This bell was simple, without an air supply from the surface. Diving bells ventilated with air from the surface were used only from the 16th century, probably the first effective means of staying under water for any length of time. In the 17th and 18th centuries, numerous devices were developed to pump air from the surface to the diver, freeing the diver from the confines of the bell by means of small diving helmet, without having to fill large bells with air. In 1819, Augustus Siebe invented the 'open' diving dress, which consisted of a waist-length jacket and a metal helmet with a hose connected to a pump at the surface.

1.1.3.SCUBA (Self-Contained Underwater Breathing Apparatus)

In 1808, Friedrich von Drieberg devised a bellows-in-a-box system worn on a diver's back and supplied with compressed air from the surface. It did not work until 1830, when Charles Condert, successfully dived to 6 meters of seawater (msw) (1.6 ata). These units were not truly self-contained because air was supplied from the surface. There is some controversy over the issue of the first SCUBA unit which emerged in the 20th century. The excessive waste of air involved in ventilating helmets from the surface gave the inventors the idea of compressing the air in tanks that could be carried by the diver. In 1943, Jacques-Yves Cousteau and Emile Gagnan redesigned a gas-generator engine regulator, producing a device that automatically provided compressed air to a diver at the pressure determined by depth, on his slightest attempt to inhale (before this, all 'self-contained' units used a continuous supply of air, or had to be manually turned on and off). Cousteau and Gagnan attached their new 'demand valve' regulator to hoses, a mouthpiece and a pair of compressed air tanks. In January 1943, Cousteau tested the unit in the River Marne outside Paris. After a modification (placing the intake and exhaust valves at the same level), they patented the 'Aqua Lung'. The aqua lung gave a boost to the diving industry, made underwater diving simple and brought compressed air diving to the level of recreational sport.

Diving development in the following decades not only dealt with the improvement of diving gear, but also with the development of alternative gas mixtures. The advantages of air as a diving gas are its ready availability, inexpensive means of compressing and storage in cylinders, and its safe use for shallow dives (Hamilton, 1997; Joiner, 2001b). Compressed air is not the ideal gas mixture for diving, because of the adverse effect of nitrogen narcosis at depths greater than 30 msw, and the decompression liability it imposes after prolonged diving between depths of 12 and 30 msw (Joiner, 2001b). The limitations of air led to the development of gas mixtures containing less nitrogen, in order to prevent or reduce nitrogen narcosis and decompression sickness. The most common alternative mixtures used today by recreational as well as by professional divers for shallow diving (down to 30 to 40 meters), are mixtures of nitrogen and oxygen which are oxygen enriched/nitrogen depleted relative to air, collectively known as nitrox.

1.2. Patho-physiological aspects

This section will describe diving/pressure-related illnesses, such as nitrogen narcosis, decompression sickness and oxygen toxicity, and will briefly outline their physiological basis. It will then focus on nitrox mixtures and on the rationale behind their alleged reduced narcotic potency.

1.2.1. Nitrogen narcosis

A major limitation of the breathing of compressed air has come to be known as nitrogen narcosis. Nitrogen narcosis was first described, not in connection with diving, but rather in caisson work (underwater construction work in an atmosphere of compressed air which prevents water from leaking into the working area). In 1835 a Frenchman, Junod, noted that when one breathes compressed air "the function of the brain is activated, imagination is lively, thoughts have a peculiar charm and in some persons, symptoms of intoxication are present" (quoted in Bennett, 1997). Soon after, nitrogen narcosis was encountered and acknowledged during the first Royal Navy deep air-diving trials to 91 msw, and was described as a "slowing of cerebration" or "as if … under an anesthetic", but was at that time attributed to "mental instability" of some candidates (Doolette, 2003). It was later recognized to be a universal phenomenon of deep diving, and was termed: 'rapture of the depths'.

Compressed air narcosis is now recognized as being due to the increased partial pressure of nitrogen. Nitrogen is the primary component (79 % by volume) of air. It is inert, in the sense of not taking part in energy transformations and other cellular reactions; therefore, it accumulates in physical solution in tissues through the effect of Henry's Law, that states that any solvent will dissolve a quantity of gas that is proportional to the partial pressure of the gas and to the physicochemical affinity between its molecules and those of the gas, such affinity being expressed by the solubility coefficient of the gas in the liquid medium. In order to experience a narcotic effect, the gas (or for that matter, any narcotic or anesthetic agent) must dissolve into the lipid neuronal membrane at some critical amount. Every gas has its own solubility in lipid. Highly lipid-soluble gases (such as nitrous oxide) or volatile liquids (ether, chloroform) will cause narcosis at atmospheric pressure. Other gases, like nitrogen, will require elevated pressures.

Dalton's Law of partial pressures states that the total pressure of a gas mixture is equal to the sum of the partial pressures of the gaseous components of the mixture. Therefore, as the total gas pressure in the lungs increases with increasing dive depth, the partial pressure of nitrogen increases proportionally, and more nitrogen becomes dissolved in the blood and tissues. The high nitrogen concentration in the neuronal membrane impairs the conduction of nerve impulses. At first, the gas mimics the effects of alcohol or narcotics, leading to the feeling of euphoria that intensifies as the total hydrostatic pressure (depth) increases (Bennett and Rostein, 2003). Narcosis usually starts at 30 msw, and appears as a mild impairment of performance in unpracticed tasks. Mild euphoria might even cause delayed responses to visual and auditory stimuli (Table 1). The impairment of decision-making ability can be more clearly defined as causing one to become 'drunk', as with alcohol (Monteiro *et al*., 1995). Monteiro *et al.* assessed subjective feeling after consuming alcohol and after a simulated 50 meter dive for 30 minutes. The preliminary results raised the hypothesis that ethanol and nitrogen act in the same way on the brain. At higher partial pressures, actual narcosis ensues, that may progress to full general anesthesia. The signs and symptoms of nitrogen narcosis in a dry atmosphere are shown in Table 1, taken from the National Oceanic and Atmospheric Administration (NOAA) Diving Manual (Joiner, 2001a). It should be stressed that the greatest danger of nitrogen narcosis is not reaching extreme symptomatology. That would be a rare event, such as a very negatively buoyant diver undergoing uncontrolled descent. Rather, the stage of euphoria, overconfidence and loss of judgment may make the diver lose alertness, take extra risks, and start a chain of events culminating in death from other causes (Bennett, 1988). One also needs to keep in mind that the safety of the dive relies on a 'buddy system', and that not being in full control and faculties jeopardizes the buddy, who may also be suffering from narcosis.

Table 1: Hyperbaric influences on cognitive and psychomotor functioning related to nitrogen narcosis with compressed air diving according to depth (Joiner, 2001a)

Depth (m)	Effect
0-30.5	Mild impairment of performance on unpracticed tasks. Mild euphoria.
30.5	Reasoning and immediate memory affected more than motor coordination and choice reactions. Delayed response to visual and auditory stimuli.
30.5-50.3	Laughter and loquacity may be overcome by self-control. Idea fixation and overconfidence. Calculation errors.
50.3-70.1	Convivial group atmosphere. Possible terror reaction in some. Talkative. Dizziness reported occasionally. Uncontrolled laughter approaching hysteria in some.
70.1	Severe impairment of intellectual performance. Manual dexterity less affected.
70.1-91.5	Gross delay in response to stimuli. Diminished concentration. Mental confusion. Increased auditory sensitivity, i.e. sounds seem louder.
91.5	Stupefaction. Severe impairment of practical activity and judgment. Mental abnormalities and memory defects. Deterioration in handwriting, euphoria, hyper-excitability. Almost total loss of intellectual and perceptive faculties.
91.5	Hallucinations (similar to those caused by hallucinogenic drugs rather than alcohol)

The lipid-solubilities of nitrogen and oxygen are 0.067 and 0.11 mol/l, respectively, meaning that physically, oxygen should be twice as narcotic as nitrogen (Bennett, 1997). The prevalent assumption of the negligible narcotic effect of oxygen is based on the fact that it is consumed by the tissues. However, its local concentration at a given site is a balance of supply and metabolic demand, and at high partial pressures, in a tissue with a relatively low oxygen consumption rate, its possible contribution to the total narcotic effect of a gas mixture cannot be dismissed.

1.2.2. Decompression sickness

While nitrogen narcosis is an affliction occurring during the bottom phase of the dive, decompression sickness (DCS) is another effect of increased partial pressure and tissue solubility of nitrogen, but one that is manifested upon ascent from depth. It is a condition of nitrogen bubble formation (Kindwall, 1995), which occurs when gas previously dissolved in body fluids under pressure evolves from solution, and forms bubbles or layers of gas in the tissues or bloodstream when pressure is reduced. It can result from a diver staying too long at depth while excessively loading the tissues with nitrogen, and then ascending rapidly to the surface, or it can occur even with a perfectly normal dive which on another occasion or in another diver did not cause symptoms (Kindwall, 1995). It should be noted that the presence of inert gas bubbles or gas phase does not necessarily in itself lead to problems: what is important is the eventual size of these bubbles, their location, and their residence time, determined by the ability or otherwise of the body to rid itself of them before they cause damage. The solution for DCS is diving according to a diving table, which defines all depth-time combinations that may be undertaken safely with a direct ascent to the surface, and by staging the ascent. This should decrease, in most divers, the formation of bubbles and their size and keep them below the symptomatic threshold.

1.2.3. Oxygen toxicity

The best way of avoiding both nitrogen narcosis and DCS is not to breathe nitrogen at all. This again relies on the fact that oxygen is a gas consumed by body cells, and as such, should accumulate less or not at all in the tissues during the dive, thus reducing the potential to form a gas phase. In addition, if a gas phase does form, it will be short-lived for the same reason, and should not cause DCS. So what will happen if one breathes pure oxygen instead of air? This has been attempted ever since the advent of compressed air dives in the late 80s and early 90s, and soon gained popularity for combat diving, since oxygen can be breathed in a closed circuit with a CO_2 absorber, thus saving gas and not emitting bubbles. Pure oxygen breathing, however, has its own severe limitations, since as essential as oxygen is for life, too much oxygen is a poison that may injure the tissues. This injury, termed oxygen toxicity, is expressed in two ways (Rutkowski, 1996):

- **Lung (pulmonary) oxygen toxicity** – Prolonged breathing of oxygen at shallow depths and even at the surface (1 absolute atmosphere or ata), will eventually cause acute pneumonia and reduce the ability of the lungs to transport oxygen to the blood. Pulmonary oxygen toxicity is not a problem with recreational divers (unless engaged in multi-day and multi-dives/day diving) because the damage usually takes many hours to occur, and divers usually do not spend enough time at depth.

- **Central Nervous System (CNS) toxicity** – CNS is the most feared hazard for divers, its advent being faster and its effects more serious the higher the partial pressure of oxygen in the breathing mix. Visual disturbances, dizziness, ringing in the ears, mood swings, and worst of all, convulsions and coma, are some of the signs and symptoms associated with CNS oxygen toxicity. The oxygen toxicity tolerance is dependent both on the oxygen partial pressure and on the length of the dive time. In order to avoid CNS oxygen toxicity while diving, the oxygen partial pressure must not exceed 1.6 ata (depth of 6 meters on pure oxygen). Recreational divers are recommended not to exceed the 1.4 ata threshold. Obviously, at such shallow depths there is no narcotic effect.

1.2.4. Nitrox mixtures

Nitrox is any mixture of nitrogen and oxygen. Since air consists primarily of nitrogen (79 vol.%) and oxygen (21 vol.%) (Although air also includes trace gases, they exist in such small concentrations that for all practical purposes they can be ignored), the surrounding air may also be termed nitrox. It has been established above that both oxygen and nitrogen have harmful effects. Yet, air may be enriched with oxygen (or its nitrogen content reduced), as long as the CNS oxygen toxicity threshold is not exceeded at the maximum depth. This is the rationale for the use of EAN in diving. In order to minimize the deleterious effects of both gases, the percentage of each gas in the breathing mixture must be adjusted according to depth and time. Table 2 displays the time a diver can spend under water without causing nitrogen narcosis or oxygen toxicity at depth, or decompression sickness upon direct ascent to the surface, according to nitrox composition and depth.

Table 2: No-Stop Dive Times (minutes) for various depths and nitrox mixtures (Joiner, 2001b)

Depth (m)	Air	Nitrox$_{32}$	Nitrox$_{36}$
15	100	200	310
18	60	100	100
21	50	60	100
25	40	50	60
28	30	40	50
31	25	30	40
34	20	30	30
37	15	25	-
40	10	20	-

It is clear that at a depth of 30 msw, raising the oxygen percentage extends the time a diver can stay underwater. However, Table 2 CNS oxygen toxicity poses absolute depth limits. That is why it is forbidden to dive deeper than 34 msw with nitrox$_{36}$. For each mixture and depth, the partial pressure of oxygen must be calculated, and the toxic threshold must not be exceeded.

The formula for calculating partial pressure (Dalton's Law) is as follows:

Partial pressure of gas = Fraction (*F*) of gas in mixture x ambient pressure

$$PO_2 = FO_2 \times P$$

For nitrox$_{40}$ at 40 msw, the FO_2 is 0.4, and the pressure is five ata:

$$PO_2 = 0.4 \times 5 = 2ata$$

The partial pressure of the oxygen is two ata, which is much higher than 1.6 ata CNS toxicity thresholds, and will result in oxygen toxicity and risk of drowning. That is why the maximum operating depth (MOD) for each proposed EAN mixture must be checked.

MOD is calculated by the formula:

$$\frac{PO_2}{FO_2} = P_{MOD}$$

The result is in ata and needs to be converted into meters. Thus, the MOD for nitrox$_{40}$ is calculated to be 1.4 ata, which is the recommended limit, and 1.6, which is the maximum limit:

$$\frac{1.4}{0.4} = 3.5ata \qquad \frac{1.6}{0.4} = 4ata$$

The results are 25 msw and 30 msw, respectively.

In order to calculate the 'no decompression' values in this table, values already established for air are redefined as Equivalent Air Depth (EAD). The EAD is the depth of a dive when breathing air that has the same partial pressure of nitrogen as of the breathing gas in question, when breathed at a given depth. The formula for EAD is as follows:

$$EAD = (FractionN_2 \times \frac{(\frac{Depth(msw)}{10} + 1)}{0.79} - 1) \times 10$$

Example: EAD for a dive with nitrox$_{40}$ to the depth of 40 msw.

$$EAD = (0.6 \times \frac{(\frac{40}{10} + 1)}{0.79} - 1) \times 10$$

$$= (0.6 \times \frac{(4+1)}{0.79} - 1) \times 10$$

$$= (0.6 \times \frac{5}{0.79} - 1) \times 10 = (0.6 \times 6.32 - 1) \times 10$$

$$= (3.79 - 1) \times 10 = 2.79 \times 10 = 27.9msw$$

This result implies that with this mixture at 40 msw, the diver may spend time as if he were diving at 28 msw on air and then surface at once. It also suggests that if it is assumed that oxygen does not contribute to narcosis, the diver will experience the same narcotic effect as a dive on air to 28 msw.

1.2.5. Lessening of narcosis by nitrox?

While no one will contest the diving-time extension afforded by nitrox, there is some dispute regarding reduced nitrogen narcosis as another benefit of nitrox diving. The U.S. Navy (Anon, 1999) agrees with this assumption, unlike Bennett and Rostein (2003), who claim that oxygen might be as narcotic as nitrogen, and together with nitrogen, it could be even more narcotic. There are some who ignore the question (Hamilton, 1997), or use statements such as "slightly reduced nitrogen narcosis" (Rutkowski, 1996).

The issue of even mild nitrogen narcosis should not be overlooked, since it may also interfere with obtaining scientific data while diving. For example, the measurement of the dimensions of an ancient ship, lying at 30 msw or more, may be inaccurate because of the influence of nitrogen narcosis. This indeed occurred in the Uluburun shipwreck excavation at a depth of 40 msw (Pulak, 1998). After using air, and finding their measurements inaccurate, helium and oxygen mixtures were used in order to obviate the effect of narcosis and provide better results.

Oxygen is twice as soluble in lipids as nitrogen (Wilmshurst, 1998), and if it is not metabolized fast enough, its narcotic effect cannot be neglected. Therefore, the question of whether nitrox should be regarded as a total or partial 'cure' for nitrogen narcosis remains open, and because of its great significance, it should be put to the test. On a theoretical basis, presuming nitrogen to be the sole cause of nitrogen narcosis, an air dive to 23 msw could be replaced with a 30 msw dive on nitrox$_{36}$, both dives giving an equivalent 'narcotic effect' (Edmonds, 2002). Strict experimental verification for this assumption has been sought but never presented (Bennett and Ackles, 1970; Linnarsson *et al.*, 1990; Edmonds, 2002). Dealing with nitrogen narcosis is by practicing safe dives, i.e. do not exceed 30 msw, and never exceed 50 msw (Lowry, 2002). Some research suggests that there is some adaptation to narcosis with daily diving (Lowry, 2002).

1.2.6. Summary

To summarize the physiological overview, it would seem that the human body is not designed to withstand high-pressure gases, that diving with compressed gases is very hazardous, and that perhaps it should not practiced. In reality, the accident record relative to dive hours shows this sport to be very safe, and the quest for knowledge and improvement of understanding can make it even safer. This research will seek a better understanding of nitrogen narcosis in order to help develop tools for reducing and dealing with it. For logistical reasons, it will concern depths relevant to the wide sports-diving community, at which patho-physiological symptoms of nitrogen narcosis are quite subtle, as seen in Table 1, and the psychological realm, concentrating on cognitive impairment.

1.3. Psychological aspects

Although narcosis has objective physiological attributes that may be assessed experimentally, such testing is rather complicated and costly. In addition, the actual impairments in cognitive performance, judgment and decision-making are important for the diver. Finally, the aim is to perform the research at moderate depths, in the realm of sport and scientific diving, depths at which effects are expected to be best revealed by sensitive mental tasks.

1.3.1. Nitrogen narcosis and performance deterioration

Dozens of experiments have been conducted in order to determine whether performance is impaired in a hyperbaric environment. The experiments addressed different skills, such as: manual dexterity – where participants were required to move ball-bearings from one place to another using a fine forceps; word fluency – where participants were required to verbally respond to a presented word with as many associated words as possible in a time frame of 60 seconds; arithmetic skill – where participants were required to solve multiplication and addition problems; and logical reasoning skills – where participants were examined for sentence comprehension and concept understanding. Memory skills were also tested by word memorizing and recall (Nevo and Breitstein 1999).

Nevo and Breitstein (1999) summed up 30 studies which tested performance deficits in the high-pressure environment (e.g. Adolfson, 1965; Adolfson and Muren, 1965; Baddeley, 1966; Baddeley *et al*., 1968; Banks *et al*., 1979; Bennett *et al*., 1969; Biersner *et al*., 1978; Davis *et al*., 1972; Fowler, 1973; Kiesling and Maag, 1962; Mears and Clearly, 1980; Philip *et al*., 1989; Phillips, 1984; Vernes and Darragh, 1982). Table 3 is a brief summary of impairment included in that report (Nevo and Breitstein, 1999).

Table 3: Summary of results: Effects of hyperbaric air pressure on cognitive and psychomotor functioning (presented as percentage of impairment relative to 1 ata) (Nevo and Breitstein, 1999, p. 45)

Pressure (ata)	**Arithmetic skills**	**Manual dexterity**	**Memory**	**Logical Reasoning**
1-3	0-5	5-15	0-5	0-5
4-6	5-10	15-25	5-10	5-30
7-10	10-20	25-35	10-20	?
11	20-40	35-45	20-30	?

The above results are all from participants tested in a dry hyperbaric chamber. Only a few experiments were performed under water, such as the one by Baddeley and Idzikowski (1985) that examined the water effect on divers: they summed up their research by claiming that at depths down to 30 meters, more than half of the impairment of cognitive functioning while diving stems from the water effect. Godden and Baddeley (1975) also conducted an experiment which tested the effects of water immersion per se on cognition. In this experiment, they had their participants memorize lists of words on the beach and at 5 meters under water, and then recall them either in the same or in the other environment. The results showed there was no difference in performance between the two environments. There was, however, a very clear context-dependency effect, meaning that if they memorized in one environment and recalled in the other, the participants remembered about 40% less than if learning and recalling occurred in the same environment. Davis et al. (1972), in two experiments in British waters, tested 16 divers twice underwater, at 3 meters and at 30 meters depth. The divers performed four tasks: a sentence comprehension test, a memory test, an arithmetic test and a manual dexterity test. In the memory test, participants were required to memorize a list of 10 nouns. This was done 4 times on each dive. On each occasion, the divers performed all the battery of tests. In the memory test, 30 seconds were allowed for encoding and 45 seconds for recalling the list and writing it down. The second and third memory trials were separated by the 30-second arithmetic test. All but the memory test showed a significant drop in performance at depth. The manual dexterity test was followed by an anxiety assessment. Anxiety was assessed by high plasma cortisol and a high adrenalin/non-adrenalin urinary excretion ratio. The result showed a clear connection between anxiety and deterioration of manual dexterity.

1.3.2. Memory functions

Memory performs in three stages: encoding, storage and retrieval (Melton, 1963). Encoding is the process of putting the data into the memory. Storage is maintaining the data until requested. Retrieval is the ability to recall the data from memory. For example, when a person is introduced to a stranger and told that "her name is Jane Doe", and on the next meeting the person states, "your name is Jane Doe". The initial introduction to Jane Doe was the encoding stage. The name was retained for later use – the storage stage. The next meeting brought up the retrieval stage when recognition occurred.

There are several variations of retrieval, among which are the recognition and the free recall possibilities. In a recognition test, the participant is presented with an item and asked whether he or she had seen this particular item before. The item itself is an excellent retrieval cue for the memory of that item. In contrast, in a free recall test, the participant has to produce the memorized items using minimal retrieval cues. The retrieval cues in a recognition test are more useful than those in a free recall. This is why performance in a recognition test (such as multiple-choice tests) is usually better than free recall test (such as essay exams) (Tulving, 1974).

1.3.3. Cognition and metacognition measurements

As regards narcosis, this research attempts to study the effects of moderately deep diving on performance and on self-assessment of performance by examining cognitive and metacognitive traits, respectively, as affected by depth, by the nature of the breathing mixture, or by both. Flavell (1976) defines metacognition as "one's knowledge concerning one's own cognitive processes and products or anything related to them", meaning the ability of one to assess and judge one's own knowledge and cognitive performance.

Metacognition involves two main functions: monitoring and control (e.g. Koriat and Goldsmith 1996; 1998). Monitoring relates to cognitive action tracking, for example, the way one judges one's memory of the information to be learned or memorized. Control relates to the manner by which the metacognitive system directs the cognitive system (Hart, 1965; Brown and McNeil, 1966; Koriat and Goldsmith 1996; 1998). This study only concerns the monitoring aspect of metacognition, although obviously the control aspect is of prime importance in decision-making, reaction to unforeseen events, etc., during diving, as well as during any other activity.

Monitoring can be assessed at three different stages:

- **Judgment of learning (JOL) –** JOL is the judgment made by the participant at the end of a learning trial regarding the likelihood of remembering the acquired information on a subsequent memory test (e.g. Koriat, 1997; Koriat *et al*. 2002; Sheffer, 2003).

- **Feeling of knowing (FOK) –** One may fail to recall an item from memory, but still feel that it will be recognized in a later test. This gradable ability to assess recognition is called FOK (e.g. Hart, 1965; Nelson *et al*, 1990; Koriat, 1995).

- **Confidence –** After the participant has answered the question, he/she is asked to judge certainty of answering. This judgment is called confidence (e.g. Koriat and Goldsmith 1996; 1998).

There are two ways of quantifying monitoring: one is by allowing participants to use an ordinal scale, which defines a description of confidence by a number (e.g. Nelson, 1988; Thompson and Mason, 1996). For example, “very low confidence (1)” up to “very high confidence (6)”. This is known as 'ordinal judgment' (Glenberg and Epstein, 1987). A second way to monitor success is by using a probability scale with a maximum of 100 and a minimum that depends on the number of alternatives. This method is called “assessed probability” (e.g. Lichtenstein *et al*., 1982; Keren, 1988; Koriat and Goldsmith, 1996).

In order to match monitoring with actual performance, two other measurements may be taken into consideration: calibration and resolution.

Calibration is the absolute compatibility between subjective assessments and the actual correctness of answers. It is expressed as the difference between the average of the assessments and the proportion of the correct answer. This difference is labeled as 'bias score' (e.g. Lichtenstein *et al*., 1982; Nelson *et al*., 1990). Over-confidence will occur if the subjective assessments exceed the correctness score; and vice versa, under-confidence will occur if the subjective assessments underestimate the actual correct answers (Lichtenstein *et al*., 1982; Koriat and Goldsmith, 1996).

Resolution is the ability to distinguish known items from unknown items (Nelson, 1984; Yaniv *et al*., 1991; Koriat and Goldsmith, 1996; Koriat, 1997). Resolution may be expressed by two different measurements: gamma-correlation and discrimination.

- Gamma-Correlation. The most common measure of resolution in metacognition is within-person gamma-correlation (Nelson, 1984). Ranging between -1 and +1, it measures the association between JOL/confidence and accuracy. A high positive gamma-correlation value implies a high percentage of questions for which the chance of answering correctly (or the fact of so doing) has been rightly assessed and scored (e.g. Nelson, 1984).

- Discrimination. This measure exercises a discrimination index, ranging from -100 to +100, consisting of the difference between mean monitoring judgments of correct and incorrect responses. The higher the mean JOL or confidence for the right answers than for the wrong answers, the better is the discrimination (e.g. Schraw *et al*., 1995).

Over and above, defining group/universal effects of depth and breathing mixture, the above three measurements may also be used as a screening tool to distinguish between most and least impaired individuals.

A cognition and metacognition test was conducted on Mount Everest by Nelson *et al*. (1990). There are many examples of bad decisions being made at extreme altitude. In terms of memory, learning of new information becomes impaired (Nelson *et al*., 1990). The use of metacognition helped to measure effects of extreme altitude on people’s judgment about whether or not they could retrieve information from memory. In order to examine metacognition a battery of tests was devised, constructed of general information facts, examining long-term memory and people's metacognition about their retrieval and feeling of knowing. The participants were asked at different heights in order to display a change in measure. The results of Nelson *et al*, (1990) showed a lack of effect at extreme altitude on retrieval of general information from memory. However, it was found that higher altitude presents a decline in the feeling of knowing (FOK) in the absence of any decline in retrieval. The high altitude research of Nelson *et al*. (1990) inspired this research, which uses the same metacognitive tools for an underwater experiment.

1.3.4. Consistency in metacognitive accuracy

There are large individual differences in the degree of association between the accuracy of memories and subjective confidence in those memories. Are these differences stable within the same test, and between alternative forms of a test? Thompson and Mason (1996) tested consistency using three experiments and found no consistency in the accuracy of judgments in various recognition memory tasks. In contrast, Sheffer (2003) did find consistency between the accuracy of resolution scores in different cognitive domains (verbal, quantitative and perceptual), using problem-solving items modified and adapted from out-of-date psychometric entrance tests to Israeli universities (PET). These data indicate the need for the development of procedures that will produce stable estimates of individuals' metacognitive accuracy. Finding within-participant consistency in calibration and resolution is the main goal in accepting metacognition as a reliable parameter. Knowing that the experiments are stable will help in developing an individual bar for a screening tool.

1.3.5. Summary

The physiological and psychological aspects of nitrogen narcosis have been examined, as outlined above. Most studies concerning nitrogen narcosis and its effects on performance were performed in a hyperbaric chamber, which by offering a dry and well-controlled environment, allowed the study of the more extreme effects of nitrogen narcosis, while ignoring the effects of submersion. Various experiments with psycho-motor functioning, such as arithmetic skills, manual dexterity memory and logical reasoning, all resulted in a decrease of performance. Metacognition had not been evaluated in the hyperbaric environment, but judging by the studies at high altitude (Nelson *et al*., 1990), such evaluation is very relevant and necessary.

1.4. Objectives

This research attempts to construct an effective battery of memory tests, which will be able to achieve the research goals during a short, moderately deep dive. The main goal is to assess the effect of nitrogen narcosis on a diver's memory and metamemory performance. Research shows that success or failure in performance is influenced not only by one's abilities and limitations but also by one's judgment. Nowadays it is widely acknowledged that one tends to act according to one's own judgment, whether or not it is the right thing to do (Koriat and Goldsmith, 1996; Koriat, 1997; Sheffer, 2003). Nelson *et al*. (1990) quote Hornbein, who wrote that during his climb on Mount Everest he was concerned "whether I'd remember how to blow up my air mattress at 26,000 feet" (Hornbein, 1965, p.32). Hornbein, a climber and a scientist, describes self-awareness of a problem that might happen at extreme altitudes; the problem being whether he indeed remembers a routine and **enrooted** action, and even more important, whether he acts accordingly and performs the correct action. The impairment in performance due to hypoxia at high altitude is in many ways similar to that of nitrogen narcosis in the underwater environment. When a diver intends to perform a mission under water, the ability to judge performance or limitations accurately can have a crucial effect both on the mission and on possible risk-taking. A wrong judgment might lead to a flawed performance, easily leading to harm or even death. This is why in the same fashion, as Nelson *et al*. (1990), who took metacognition out of the testing room and up Mount Everest, this research would be the first to use metacognition measurements under water. Using underwater cognition and metacognition measurements helps to achieve a better understanding of behavior under the effect of nitrogen narcosis in the actual field conditions of the open sea. Nitrogen narcosis may adversely affect the quality of judgment. This aspect, which was given ample qualitative support has not yet been systematically explored, and is a main theme of the present work.

The study's main objective was to verify or refute the alleged reduced narcosis of nitrox_{36} over compressed air. Secondary objectives were finding out whether narcosis affected the encoding or the recall phase of new information, or both, quantifying the often described over-confidence induced by nitrogen narcosis, and seeing whether it is affected by the breathing mixture; and finally, examining the consistency over time of nitrogen narcosis effects on the individual diver. The existence of such consistency implies that nitrogen narcosis affects different divers to different degrees in an inherent, time-invariant manner. This knowledge may allow the construction of a screening tool for nitrogen-narcosis-sensitive divers.

CHAPTER 2: METHODOLOGY

2.1. General

The main goal of this experiment was to compare the narcotic effects of air and nitrox on divers. Cognitive and metacognitive measurements were used to try to detect a mild nitrogen narcosis effect in the moderate experimental depth of 30 msw. The two experimental breathing mixtures at depth were air (30A) and nitrox$_{36}$ (30N). The control condition, (which incorporated all the elements of diving except the depth) was a 2 msw dive with air (2A). The experimental design was self-controlled (each participant serving as his/her own control and undergoing all types of dives on different occasions) and double-blind (neither the diver nor the experimenter were aware of the nature of the mixture breathed on any given deep dive).

The means of obtaining the cognitive and metacognitive measurements was a battery of three tests administered to the participants before, during, and following their dives. Each test included an encoding phase, followed after an interval by the actual task of recall. This was accompanied by each participant both predicting and evaluating performance *post-factum*. In order to try to separate narcotic effects on the encoding and the recall, the tests conformed to three different experimental combinations: encoding on the surface and recall at depth, both phases at depth, and encoding at depth and recall on the surface. The performance scoring and the metacognition scaling were expected to provide both individual and group quantification of the narcotic effects on the participants. Since it was unclear which type of cognitive test would work best (would be most influenced by gas narcosis) or would work at all conditions, different tests (described below) were used for each of the above three conditions. This experiment was approved by the University of Haifa's Ethics Committee (Helsinki Committee).

2.2. Experimental site

The experiments were carried out at the Israeli Maritime College in Mikhmoret, at two flat-bottomed locations chosen to provide the desired depths: the marina for the 2 msw (inshore) dives, which examined the effect of submergence per se, and a buoy anchored 2.4 km off shore, for the 30 msw dives. The experiments were conducted between December and February, the end of fall to winter. Due to the long period, the water temperature decreased from 21 °C to 15 °C. However, all dives were in calm sea (no seasickness), with no currents, the visibility was 10 meters or more, and the descent was on a plummet to a sandy bottom.

2.3. Participants and experimenters

The participants in these experiments were 35 divers: 23 male and 12 female, between the ages of 20 and 30, with a diving history range of 13 to 44 dives. All were Hebrew-speaking undergraduate students at the Israeli Maritime College in Mikhmoret who were paid per dive for participating in the experiment. Prior to participating in the experiment, each participant read and signed a standard medical questionnaire (Appendix 1), an informed consent form (Appendix 2) and received a letter describing an overall view of the experiment (Appendix 3). The total number of participants for Experiment 1 was 23, and for Experiment 2 was 26.

Not all 35 participants finished all dives/tests, due to schedule or/and personal problems. Organizing the diving schedule was difficult, since all the participants were active students who were free at different hours. Accordingly, there were irregular time gaps between dives for each participant. The shortest inter-dive gap was four days and the longest was 42 days. It is important to mention that before each dive there was a detailed briefing (Appendix 4) which made certain that no participant would forget the experimental procedures.

This experiment was managed by a team of experimenters headed by the Experiment Manager. Task allocations were as follows:

- The Experiment Manager was in charge of making up the questionnaires, administering the pre- and post-dive training/tests, monitoring the diving schedule (depth, questionnaire), logging the participant number, diving time and questionnaire number. Other than designing and controlling the scientific aspects, his main role was to organize the participant, the dive type and the test in order to prevent a participant from repeating the same questionnaire on different dives.

- The Diving Supervisor was in charge of filling the cylinders, checking the dive mixtures, providing diving equipment, and keeping track of safety issues. The main responsibilities were to oversee all non-scientific and safety aspects of the dives, and to verify that participants received the right mixture for the right dive.

- The Timekeeper was in charge of timing the phases of the experiment. This was done by the timetable slate and an underwater horn. The Timekeeper was also a scuba instructor and was the experiments' Safety Diver. He was in charge of the dives and divers, and had the power to interrupt the course of the dive or to abort it whenever he deemed necessary.

- The Tester was an advanced diver (dive master or scuba instructor) and his main role was to act as a dive buddy to the participant (each participant had his own Tester). His role was also to carry the questionnaire crate and to hand out and receive the questionnaires according to the Timekeeper's signals.

- The Skipper was in charge of the boat, and his additional roles were keeping an eye out for approaching boats, safety swimmer, assisting divers in case of an emergency (oxygen administration, etc.).

The maximum number of personnel on each round was seven divers (three participants, three Testers and a Timekeeper) and a Skipper. The Timekeeper was in charge of placing the divers under water so that he was able to see them and only be seen by the Testers. The only people who knew the nature of gas mixture for any given dive were the Diving Supervisor and the Skipper.

2.4. Equipment

The equipment consisted of diving gear, experimental equipment, safety equipment and transportation. Each diver used personal diving gear which included a nitrox-compatible single hose scuba regulator, weights, buoyancy control device (BCD), diving suit, mask and fins All inshore dives were made with air in air-marked cylinders, all offshore dives used nitrox-marked cylinders, containing nitrox$_{39}$ for the Testers, and either nitrox$_{36}$ or air for the participants. All cylinders were filled and checked at the University of Haifa maritime workshop. The experimental equipment comprised a laptop, with which on-the-surface training and tests were administered and all experimental data stored, printable and writable polyester A4 paper (DFDM100) fastened on both sides of rough Teflon writing slates for the underwater testing, pencils, carrying crates, a timekeeping slate and an underwater horn. The safety equipment included a nitrox$_{50}$ cylinder and a safety-stop bridge at 5 meters depth for desorbing nitrogen, a 100% oxygen cylinder on board the boat and a diving flag. Transportation to the offshore dive site was by boat, which was also used as a gearing-up platform and safety boat.

2.5. General procedure and test description

Each participant performed under three dive conditions (2A, 30A, 30N) on separate days, in random order. Dives were usually spaced a few days apart, but some were up to a month apart. On each dive, the three types of tests were presented, and for each test, three versions were composed, one for each diving condition. Pilot studies of the test versions were examined for their balance, and proved that the three versions of each test were counter-balanced in terms of cognitive results. All tests were put in a Latin square so the various versions of tests and gas mixtures were divided equally. However, the order of the three experiments remained the same, first the Faces Test (Experiment 1), then the Word-pair Test (Experiment 2), and lastly the picture Test (Experiment 3).

In this experiment, participants breathed air and nitrox$_{36}$ at 30 msw. The maximum limit for nitrox$_{36}$ is:

$$\frac{1.4}{0.36} = 3.89ata \qquad \frac{1.6}{0.36} = 4.44ata$$

To reach a 1.4 ata oxygen partial pressure one would dive to 29 msw (or 34 msw to reach a 1.6 ata oxygen partial pressure). This equivalent air depth of nitrox$_{36}$ to the depth of 30 msw is:

$$\begin{aligned}
EAD &= (0.64 \times \frac{(\frac{30}{10}+1)}{0.79} - 1) \times 10 \\
&= (0.64 \times \frac{(3+1)}{0.79} - 1) \times 10 \\
&= (0.64 \times \frac{4}{0.79} - 1) \times 10 = (0.64 \times 5.06 - 1) \times 10 \\
&= (3.24 - 1) \times 10 = 2.24 \times 10 = 22.4msw
\end{aligned}$$

This result implies that at 30 msw with this mixture, the diver may spend time as if he was diving at 22.4 msw (35 to 40 minutes) on air, and then surface at once. However, since this research is double-blind and neither the participants nor the testers know the gas mixture, the calculations were done by calculating air .This is why bottom time did not exceed 20 minutes.

2.6. Experiment 1

The first test, Experiment 1, was the Faces Test (Appendix 5-10), in which encoding (phase one) was made on shore and recall (phase two) was performed under water. It addressed the effects of depth and mixture on the ability to recall items memorized just before the dive, and on how participants judged their performance.

2.6.1. Participants

Twenty-three divers participated, six female and 17 male. All participants were Hebrew-speaking undergraduate students at the Israeli Maritime College in Mikhmoret, who were paid per dive for participating in the experiment.

2.6.2. Test construction

Test versions were constructed from a pool of 270 different grayscale images of faces taken from various published collections on the internet. The faces were categorized according to gender, similarity and photo quality, and were put into groups of four. Each group of four faces represented a question and one of the faces was arbitrarily chosen as the correct answer and was put into the encoding list. The latter contained, in addition to the 20 faces for testing, another 210 faces for distraction. The gender of the faces was divided equally among the three versions of the test. The ratio between female faces and male faces in the questionnaire was kept at 9:11.

2.6.3. Procedure

In phase 1, the encoding phase, participants were allotted 2 minutes 30 seconds on shore for encoding 30 faces individually displayed in succession, five seconds per face, as MS-PowerPoint slides on a laptop screen (Figure 1 is an example), In addition, participants were asked to note their ability to recall this face among four alternative faces (JOL), using a scale ranging from 25%

(chance level), to 100% (absolute certainty). The recall phase began approximately 45 to 60 minutes after the encoding phase, after the diver had descended to the required depth. This phase consisted of 20 items, divided into two stages (1 minute per stage) consisting of 10 items each, broken by 10 seconds for rest and Tester preparation for the next stage. The 6 seconds time per page were synchronized by the Timekeeper.

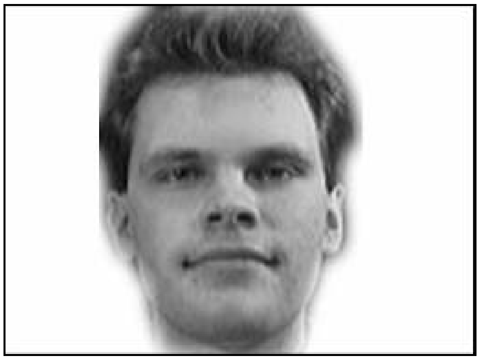


Figure 1: Faces encoding slide – an example of an encoding MS-PowerPoint slide from the Faces Test. The participant has to memorize this face for 5 seconds and note his/her JOL Time per slide was 5 seconds.

מידת הביטחון שהתשובה
נכונה___________________

Figure 2: Faces recall slide – an example of a recall slide from the Faces Test. The participant has to circle the correct face and note confidence in his/her answer. In Hebrew: the confidence in your answer is ___. Time per slide was 6 seconds.

In the recall phase, participants were presented with 20 printed pages in succession in 2 minutes, each with 4 faces, only one of which had been shown to them in the encoding phase (Figure 2 is an example). Participants were asked to mark the familiar face (target face) as well as to grade their confidence in their answer, using a scale ranging from 25% (chance level), to 100% (absolute certainty).

2.7. Experiment 2

The second test, Experiment 2, was the Word-pair Test (Appendix 11-13), both phases of which were performed under water, thus testing depth and mixture effects on both the encoding and recall aspects of memory and on the JOL.

2.7.1. Participants

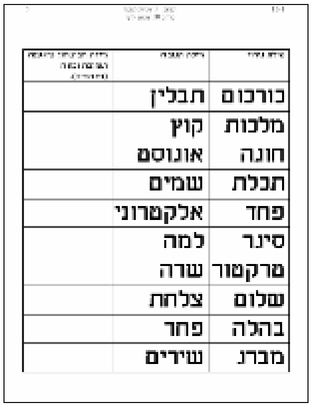
Twenty-six divers participated in Experiment 2: 8 female and 18 male. All participants were Hebrew-speaking undergraduate students at the Israeli Maritime College in Mikhmoret, who were paid per dive for participating in the experiment.

2.7.2. Test construction

The stimulus lists were constructed based on the results of a preliminary experiment designed to obtain data on the perceived relative difficulty of the items based on Hebrew word association norms (Rubinsten *et al*., 2005). For this experiment, 75 Hebrew stimulus-response word-pairs were compiled, which represented a wide range of associative relatedness. The 75 pairs of words were divided into two groups – easy and medium difficulty. An additional 150 words were paired to a third, difficult group. The three groups: easy, medium, and difficult, were organized into three different stimulus list versions, each containing 50 pairs: 10 easy, 15 medium, and 25 difficult.

2.7.3. Procedure

In phase 1, the encoding phase, participants were instructed to memorize a visual list of 50 paired associates in order that 30 seconds afterwards, they would recall the target word (response term) when presented with the cue word (stimulus term) (phase 2). For each word-pair in phase 1, they were asked to grade, on a scale of 0-100, the likelihood of recalling the response term during the test (JOL) (Figures 3 and 4 are examples). In this test, confidence was not assessed, the reason being that bottom time constraints dictated a very short lag between the two phases, too short to ensure lack of bias of the confidence grading by the JOL grading.



כורכום	תבלין	
מלכות	קוץ	
חונה	אוגוסט	
תכלת	שמים	
פחד	אלקטרוני	
סינר	למה	
טרקטור	שרה	
שלום	צלחת	
בהלה	פחר	
מברג	שירים	

Figure 3: Word-pair encoding – an example of an encoding slide from the Word-pair Test. The participant has to memorize the stimulus and the response word, and note his/her JOL in the answer after each pair. Time per slide was 50 seconds.

קרואסון	
פלאפל	
רצפה	
כבש	
כפתור	
פנקס	
שלום	
שביתה	
הסמיק	
מגעיל	

Figure 4: Word-pair recall slide – an example of a recall slide from the Word-pair Test. The participant has to write the response word according to stimulus word. Time per slide was 50 seconds.

The lists were divided into 5 pages, 10 word-pairs per page and 50 seconds (10 seconds per word-pair) allotted to memorize them and note the JOL – for a total of 4 minutes 10 seconds for phase 1. Page switches were synchronized by the Timekeeper.

Phase 2 commenced after a 30 seconds break of STM consolidation used for checking air status. During the test phase, the 50 stimulus words were represented the same way (5 pages, 10 words per page) but in a different order (of pages and of words in a page), and with adjacent blank spaces, in which the participants had to write the response word. Time per page and total time for phase 2 was the same as in phase 1.

2.8. Experiment 3

Experiment 3 was the Picture Test (Appendix 14-19). It was designed to isolate the effects of depth and breathing mixture on the encoding of visual memory, by recalling back on the surface visual information memorized underwater. In phase 1, performed during the last 30 seconds of bottom time, participants were presented with, and asked to memorize, a single multi-detailed picture, and on a scale of 0-100, score their overall prediction of how well they would be able to recount its details (JOL) (Figure 5 is an example). Then, back on shore, 30 minutes after surfacing, they were asked 20 multiple-choice questions concerning details from the picture (phase 2) in a time limit of 5 minutes (Figure 6 is an example).

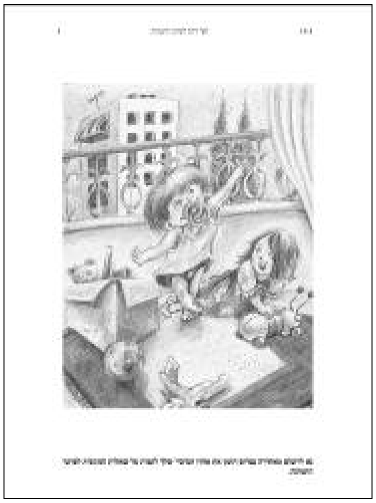
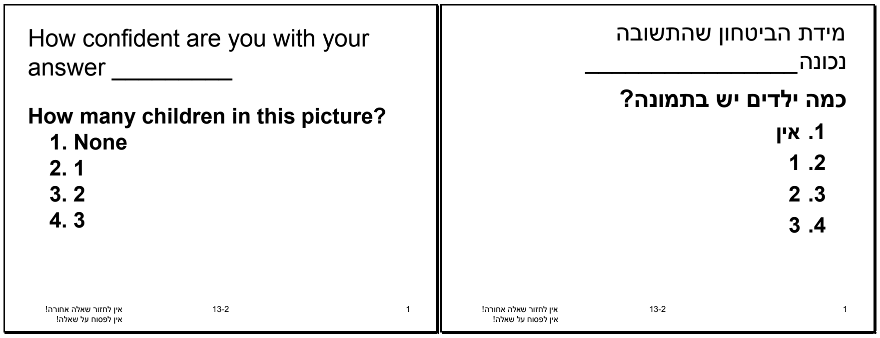


Figure 5: Picture encoding - an example of an encoding slide from the Picture Test. The participant had to memorize the details of the picture. Time per slide (+ JOL) was 30 seconds.



How confident are you with your answer _________

How many children in this picture?

1. **None**
2. **1**
3. **2**
4. **3**

אין לחזור שאלה אחורה!
אין לפסוח על שאלה!
13-2
1

מידת הביטחון שהתשובה נכונה________________

כמה ילדים יש בתמונה?

1. אין
2. 1
3. 2
4. 3

אין לחזור שאלה אחורה!
אין לפסוח על שאלה!
13-2
1

Figure 6: Picture recall - an example of a recall slide from the Picture Test. The participant had to circle the correct answer and note his/her confidence in the answer. Time limit per slide was 15 seconds. (Left figure is a translation not original)

Unfortunately, due to problems such as picture quality and ambiguous questions, it transpired that there was a clear difference between the results obtained from the three picture versions, even in the shallow dives, making it impossible to know whether a participant indeed performed badly or just had an unclear picture and/or questions. Since judgment was also dependent on picture quality/clarity, this entire test had to be abandoned.

2.9. Experimental outline

Figure 7 is a time-depth profile scheme of the dives. Bottom time and its divisions were strictly adhered to, although some leeway was permitted for the descent and positioning on the bottom. Other times are approximate. A given trial started with a Faces presentation encoding on shore, followed by time spent for organizing diving equipment and arriving at the diving site. For shallow dives, the participants walked to the near-shore spot. The latter interval was not strictly kept and lasted from 40 to 60 minutes. Figure 7 shows a 40-minute interval between the encoding to the face images and entering the water, and another five minutes for the descent - total 45 minutes.

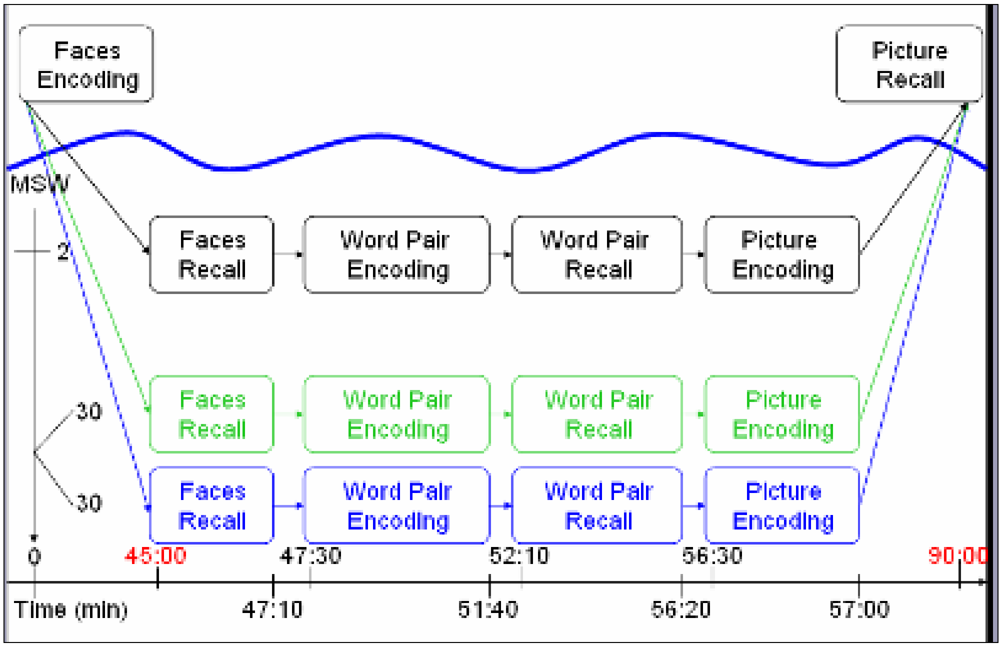


Figure 7: Experiment process on a time schedule. Black - 2 msw (shore dive) air in tanks. Green - 30 msw (deep dive) nitrox$_{36}$ in nitrox tanks. Blue - 30 msw (deep dive) air in nitrox tanks. Red - estimated time - not always accurate.

All the dives were performed with the Tester as a diving buddy. After descent, the Timekeeper organized the divers so that he stood behind the participants (usually 3), who were on their knees on the bottom, facing their Testers who were in sight of the Timekeeper (Figure 8). This position gave the Timekeeper a good view of the participants' progress and eye contact with the Testers for mutual signaling. Horn signals throughout the dives were: one blow for 'start', two blows for 'next' (new page) and three blows for 'stop'. Tapping two metal rods was used as a backup.

Figure 8: Underwater photo. From front to back: Timekeeper (back to camera), three participants (back to camera) and three Testers facing participant and Timekeeper (Photo by S. Breitstein)

As soon as everybody had settled on his or her knees, the Timekeeper gave the signal to begin the Faces recall phase, Experiment 1. The Timekeeper gave the sign to switch to the Word-pair encoding (Experiment 2 phase 1), 20 seconds after completing Experiment 1, followed by a safety check. After the check, the Timekeeper signaled to begin the Word-pair recall (phase 2). This was followed after a 10-second break by the Picture encoding (Experiment 3, phase 1), after which all divers began the ascent to a 3-minute safety stop at 5 meters, switching to nitrox$_{50}$ mixture for faster nitrogen adsorption. All underwater time measurements were strictly kept by the Timekeeper and did not exceed the 20 minutes no-decompression bottom-time limit. Participants answered questions regarding the Picture recall (phase 2) on shore and off gear, 30 minutes after beginning ascent. This procedure was followed for three dives using three versions of the tests. The time span between each dive ranged from a few days to a more than a month, there never being more than one dive a day.

2.10. Cognitive and metacognitive evaluation

The data noted by the participants were used as raw data for scoring the correctness of the answers, and for computing individual and group mean JOL and confidence.

2.10.1. Correctness

Correctness performance was measured as the percentage of the correct answers the participant answered on each test. Unanswered questions were marked as mistakes. For example, if a participant recognized correctly 15 faces out of the 20 target faces, gave wrong answers for two and did not answer on three Faces questions, the total correctness was scored as 75%.

2.10.2. Judgment of learning

Participants, during the learning phase or immediately following it, were asked to evaluate how well they had learned/memorized their task, by predicting how well they would do when asked to perform by drawing on their memory. This assessment is also known as the Judgment of Learning, or JOL measurement. For instance, while exposed to the faces, the participants had to assess on a scale of 0-100%, for each face, their ability to distinguish it when presented 45 minutes later on the screen, among three distracters. In the Faces Test, individual average of the percentage values were taken only from the 20 chosen faces, which were used as the target faces, and not from all 30 faces presented.

2.10.3. Confidence

Participants were asked during the recall phase to evaluate how well they completed their tasks. This assessment is also known as the confidence measurement. For instance, while exposed to the target face question, after marking the chosen face out of the four, the participants had to grade on a scale of 0-100%, how confident they were in their choice.

2.11. Metacognitive measurements:

The metacognitive measurements, which were derived from the correctness, JOL and confidence scores, were the calibration and resolution.

2.11.1.Calibration

The measurement of calibration is the absolute compatibility between subjective assessments and the actual correctness of the answers. The calculation is achieved by deducting the mean JOL or mean confidence from the percentage of correct recall for each participant. A zero value shows perfect judgment/evaluation, which means that the participant knows how much he/she knows. Positive calibration values represent over-confidence, while negative calibration represents under-confidence. However, calibration is only a crude parameter, since for a performance of 60%, one could get an average confidence value of 60% (zero calibration) by stating a 60% confidence for each of 20 response words, or by giving zero scores for the eight words which were wrong and 100 scores for the 12 which were correct (obviously, a better self-judgment), and all grades in between. Therefore, the measurements of resolution were devised to give a better probe into the quality of judging performance.

2.11.2.Resolution

Resolution in this work was assessed by two measurements: gamma-correlation and discrimination.

Gamma-Correlation: The within-person gamma-correlation evaluates all pair-combinations of scores and correctness for the individual questions posed to the participant. It gives a coefficient ranging between -1 and +1, with (rather rare) negative values indicating a consistent misjudgment of performance.

Discrimination: Discrimination, a cruder measure of resolution, is calculated as the difference between mean JOL or confidence in the right answers, and mean JOL, or confidence in the wrong answers. As long as an average score for the right answers is higher than the average score for the wrong answers, this demonstrates some degree of discrimination, and the higher the difference, the higher is the resolution (e.g. Schraw *et al*., 1995).

2.12. Statistics

The significance of this experiment went through several statistical phases using the SAS program and MS-Excel. The differences between the three diving conditions in cognitive and metacognitive performance were tested by one-way ANOVA (a within-participant design) and *post hoc* (Scheffe) comparisons, at a p level < 0.05. The Pearson correlation coefficient was used in order to define within participant inter-dive consistency in a participant's results.

CHAPTER 3: FINDINGS

3.1. Experiment 1

In this experiment 23 participants; 6 females and 17 males, fully completed all required tasks. As a result of the procedure outlined above, for each participant, diving condition and item, there were three scores: the correctness of the answer (1 or 0), the participant's JOL while being exposed to the faces, and the participant's confidence estimation in the correctness of the answer. Table 4 depicts the mean values of these three measurements for the three conditions of diving (2A, 30A, 30N)

Table 4: Experiment 1, the Faces Test results. Mean, standard deviation and range of cognition and metacogniton measures in percentage (N = 23).

Condition	**2A**				**30A**				**30N**			
Measurements	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max	Min
Correct	59.56	18.88	90.00	25.00	60.43	15.44	95.00	40.00	61.95	17.49	95.00	30.00
JOL	58.09	12.35	82.25	38.60	55.77	12.43	80.2	26.25	56.81	14.60	82.45	30.05
Confidence	58.90	13.21	89.65	41.56	60.43	16.21	92.75	31.11	59.97	15.52	85.35	30.30
Calibration	-0.65	12.70	23.50	-25.00	0.00	12.69	20.40	-32.90	-1.98	15.06	25.35	-28.50
Gamma- correlation	0.48	0.31	1.00	0.11	0.53	0.28	1.00	0.02	0.47	0.27	0.92	-0.22
Discrimination	21.06	15.02	-1.47	54.79	22.75	13.37	5.17	49.83	16.63	11.08	-1.86	43.07

Figure 9 presents the results given in Table 4, showing the clear resemblance between results at all three conditions. Figure 10 however depicts the perfect calibration achieved by the participants. It is clear that the calibration remained between -1.98% and 0% which means no over or underconfidence at all.

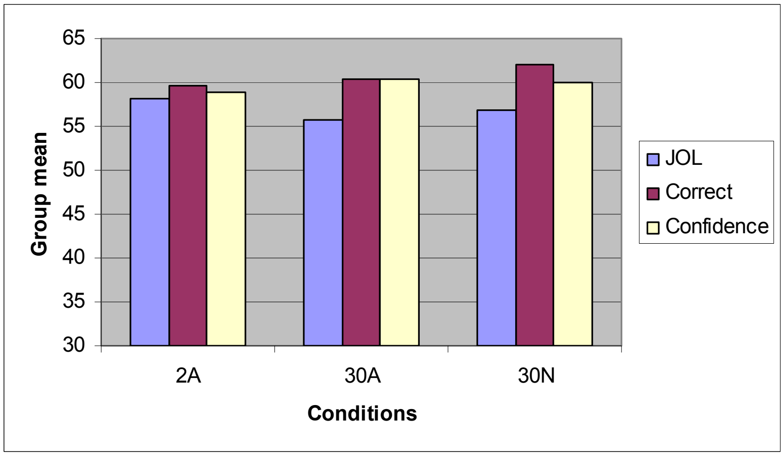


Figure 9: Experiment 1, participants' group mean JOL, correctness and confidence at the three diving conditions (N = 23).

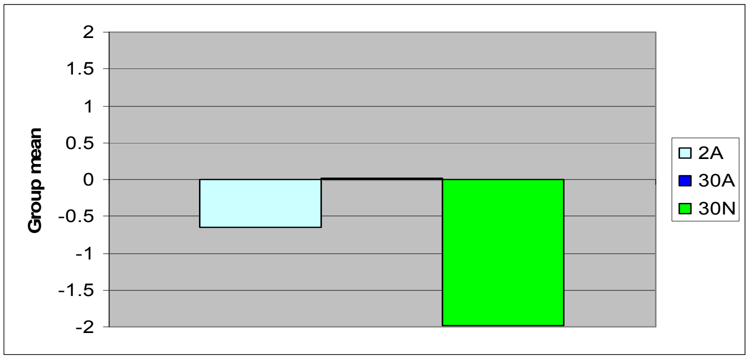


Figure 10: Experiment 1, group mean confidence (measured underwater) calibration scores as a function of the three diving conditions (N = 23).

Figure 11: Experiment 1, group mean correctness and confidence scores for nitrox vs. air for the two deep dives (N = 23).

Figure 11 shows the results of Experiment 1 in terms of correct answers and confidence, in deep diving (30 msw), breathing either air (blue), or nitrox (green). The graph reflects a slight but insignificant tendency to answer correctly when nitrox was used. There was no difference in confidence.

3.1.1. Cognitive performance

Correct answers. Cognitive performance for the Faces Test was defined as the percent of correct faces identified. The group-mean percent of correct answers (Table 4) was 59.6% for the shallow water dive, a range of 25% to 90%, 60.4% for the deep dive on air, a range of 40% to 95%, and 62% for the deep dive with nitrox$_{36}$, a range of 30% to 95%. The one-way ANOVA on diving conditions showed no significant difference between the mean scores (F<1) (see Figure 9.11).

Figure 12 shows the frequency distribution of individual correctness scores at a resolution of 10. It is difficult to discern any major difference between the three dive conditions.

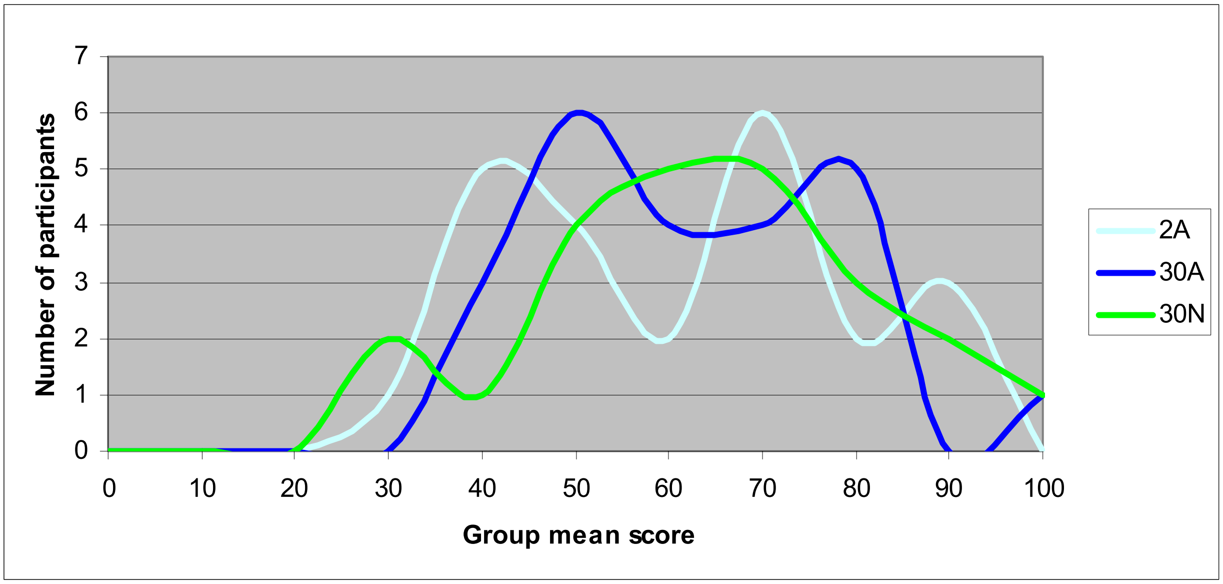


Figure 12: Experiment 1 – Frequency distribution of individual correctness scores (resolution: 10) (N = 23).

3.1.2. The metacognitive performance

Judgment of learning: JOL for the Faces Test was defined as the mean percent of JOL assessed on land while being exposed to the 30 faces. The group-mean percent of JOL (Table 4) was 58.10 for the shallow water dive, a range of 38.60 to 82.30, 55.80 for the deep dive on air, a range of 26.30 to 80.20, and 56.80 for the nitrox deep dive, a range of 30.10 to 82.50. The one-way ANOVA on diving conditions showed no significant difference between the JOL measurements (F<1) (see Figure 9).

Confidence: The measurement of confidence is the assessment of the participants of the correctness of each of their answers after it has been given. The group-mean confidence rate, shown in Table 4, was 58.90 for shallow water, a range of 41.56 to 89.65, 60.43 for the deep dive on air, a range of 31.11 to 92.75, and 59.97 for the deep dive on nitrox, a range of 30.30 to 85.35. The one-way ANOVA on diving conditions showed no significant effect on the degree of confidence (F<1) (see Figure 9, 11).

3.1.3. Derived metacognitive measurements

From the basic confidence scores derived from metacognitive and cognitive performance, metacognitive measurements were calculated: calibration, and resolution, the latter including gamma-correlation and discrimination index. The JOL for this aspect was disregarded due to the fact that it was always obtained under the same conditions: at the surface, breathing air, without the influence of depth or mixture.

Calibration: reflects the overall correspondence between confidence and performance. The group-mean calibration values, shown in Table 4, were 0.65 for shallow water, a range of -25 to 23.5, 0 for deep dive with air, a range of -32.90 to 20.40, and -1.98 for deep dive with nitrox, a range of -28.50 to 25.35. The one-way ANOVA on diving conditions showed no significant difference between the bias score measurements (F < 1) (see Figure 10).

Figure 13 shows the frequency distribution of individual calibration values, again at a resolution of 10. Here too, it is difficult to see any major difference between the three diving conditions, except for the apparent bimodality of the 30N, which is due to the curve smoothing.

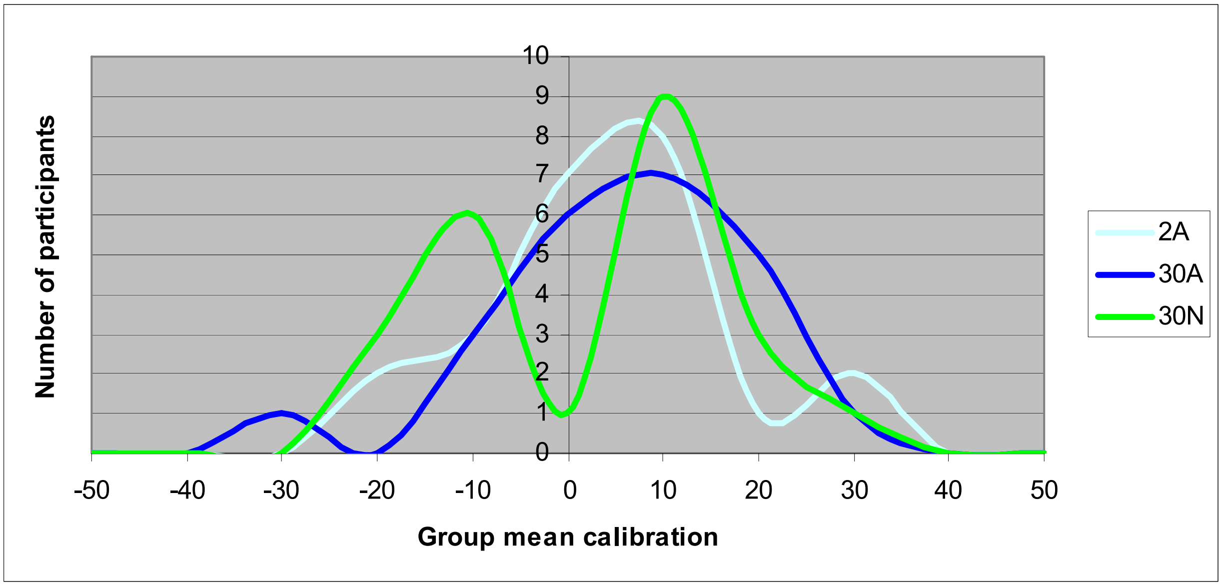


Figure 13:Experiment 1- Frequency destribution of individual confidence calibration values (resolution: 10) (N = 23).

Resolution

Gamma-Correlation: The group-mean gamma-correlation values (Table 4) were 0.48 for shallow water with a range of 0.11 to 1, 0.53 for deep dive on air, with a range of 0.02 to 1, and 0.47 for deep dive on nitrox, with a range of -0.22 to 0.92. The rather high positive value indicates that the participants were able to distinguish fairly well between what they know and do not know. The one-way ANOVA on diving conditions showed no significant difference between the gamma-correlation measures between the three diving conditions (F<1).

Discrimination index: The group-mean discrimination index (Table 4), was 21.06 for shallow water with a range of –1.4 to 54.79, 22.75 for deep dive on air, with a range of 5.17 to 49.83, and 16.63 for deep dive on nitrox, with a range of –1.86 to 43.07. One-way ANOVA between the diving conditions yielded an insignificant effect on the bias score F(2,48) = 1.53, MSE = 163.43, p < 0.23.

3.1.4. Consistency of cognitive and metacognitive performance for the three diving conditions

An examination of the correlations between conditions (2A, 30A, 30N) as regards both performance and judgment, found insignificant correlations between most combinations (Table 5). It is clear that the confidence judgments were stable, but the performance and metacognitive measures were not. The finding that the judgments are more stable than the cognitive and metacognitive performance has been found previously (e.g. Schraw *et al.*, 1995; Thompson and Mason, 1996; Sheffer, 2003).

Table 5: Experiment 1 - Participant auto-correlation coefficients between diving conditions (N = 26).

Correlation	2A-30A	2A-30N	30A-30N
Correct	.34	.21	.52
Confidence	.54*	.75**	.58*
Calibration	-.17	.41*	.30
Discrimination	.30	-.21	-.15

** $p < 0.001$ * $p < 0.05$

3.1.5. Summary

In this experiment, depth had an insignificant influence upon cognition and metacognition, as summarized in Table 4 and Figure 9. The group-mean success rate ranged between 59% and 62%, while confidence in success ranged between 59% and 61%, demonstrating excellent calibration and excellent accord between JOL and confidence. The over-confidence graph (Figure 10) is the calibration results of Experiment 1, which shows an almost complete calibration between expected and achieved performance.

3.2. Experiment 2

In this experiment, there were 26 participants: 8 females and 18 males, who completed all the tasks. As a result of the procedure outlined above, for each participant, for each diving condition and for each item, there were two scores: the correctness of the answer (1 or 0), and the participant's JOL while studying the word-pairs. Table 6 depicts the mean values of these two measurements for the three conditions of diving: 2A, 30A and 30N.

Table 6: Experiment 2 – Word-pair results. Mean, standard deviation and range of cognition and metacogniton measurements (N = 26).

Condition / Measurements	2A				30A				30N			
	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max	Min
Correct	47.15	14.69	70.00	14.00	38.69	17.19	82.00	14.00	39.00	14.93	70.00	12.00
JOL	48.66	15.98	74.22	14.14	49.49	17.50	83.72	22.10	51.04	17.38	84.08	15.66
Calibration	1.51	13.68	27.26	-28.8	10.80	13.57	43.48	-11.90	12.04	14.04	44.90	-25.46
Gamma-Correlation	0.60	0.14	0.90	0.28	0.58	0.13	0.84	0.20	0.53	0.23	0.89	0.06
Discrimination	27.79	13.56	8.37	64.29	25.91	12.51	1.54	57.52	23.02	14.80	0.79	52.62

Figure 14 demonstrates the results of Experiment 2 in terms of correct answers and JOL. In the control shallow dives (2A), the participants again demonstrated strict correspondence between their JOL and the actual percentage of correct answers. However, in both deep dives (30A and 30N) there is a gap between their JOL, which stayed the same as in the shallow dive, and the percentage of correct answers, which declined significantly. This means that the participants had lost their ability to assess themselves correctly, with the type of breathing mixture having no apparent effect.

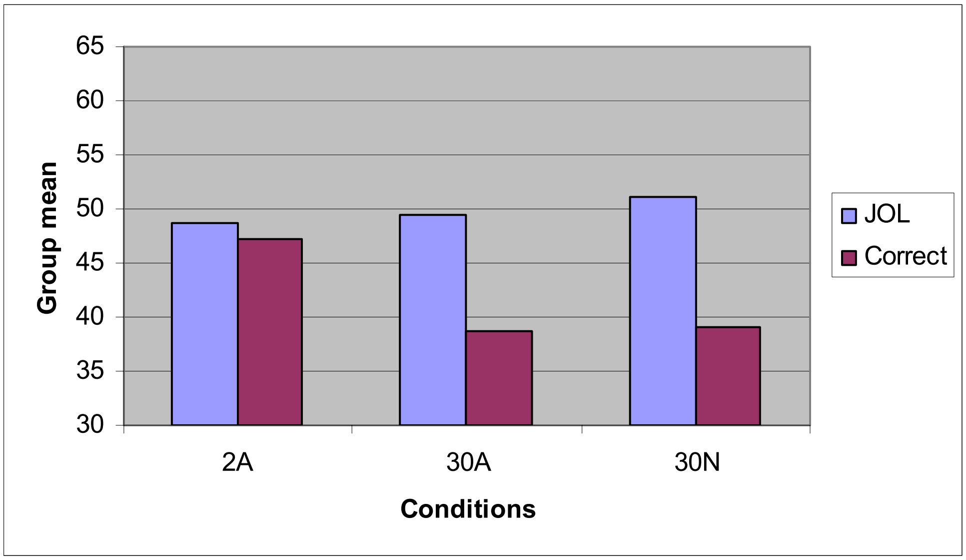


Figure 14: Experiment 2 - participants' group mean correctness and JOL at the three diving conditions (N = 26).

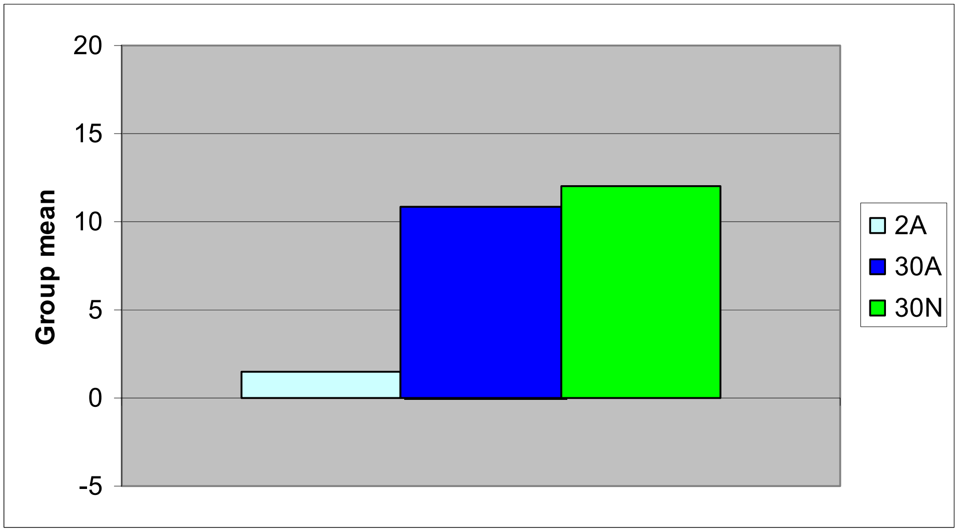


Figure 15 - Experiment 2 group mean JOL (measured under water): calibration scores as a function of the three diving conditions (N = 26)

Figure 16 - Experiment 2 group mean correctness and confidence scores for nitrox vs. air at the two deep dives (N = 26).

The calibration graph (Figure 15) is the calibration result of Experiment 2. The graph shows deterioration in calibration from shallow to deep dives. There is no significant difference between nitrox and air.

Figure 16 shows the results of Experiment 2 in terms of correct answers and JOL in deep diving (30 msw), using either air (blue) or nitrox (green). Figures 11 and 16 reveal similar results for nitrox and air at depth. The reason a comparison of confidence in the Faces Test was used (Experiment 1) and JOL in the Word-pair Test (Experiment 2) is due to the fact that the JOL in the Faces Test was performed on land, while the confidence measurement was tested under water (like the JOL of the Word-pair Test).

3.2.1. Cognitive performance

Cognitive performance for the Word-pair Test was defined as the percent of correct pairs completed. Unanswered pairs were counted as mistakes. The group-mean percent of correct answers (Table 6) was 47.15% for the shallow water dive, a range of 14% to 70%, 38.69% for the deep dive on air, a range of 14% to 82%, and 39% for the deep dive with nitrox$_{36}$, a range of 12% to 70%. One-way ANOVA between the diving conditions yielded a significant effect for the percentage of correct recall $F(2,50) = 8.71$, $MSE = 68.74$, $p < 0.001$. Post hoc comparisons (Scheffe) showed a significant difference between 2 msw and both 30 msw dives, in favor of the shallow dives, but no significant difference between the 30 msw nitrox and the air dives (see Figures 14 and 16).

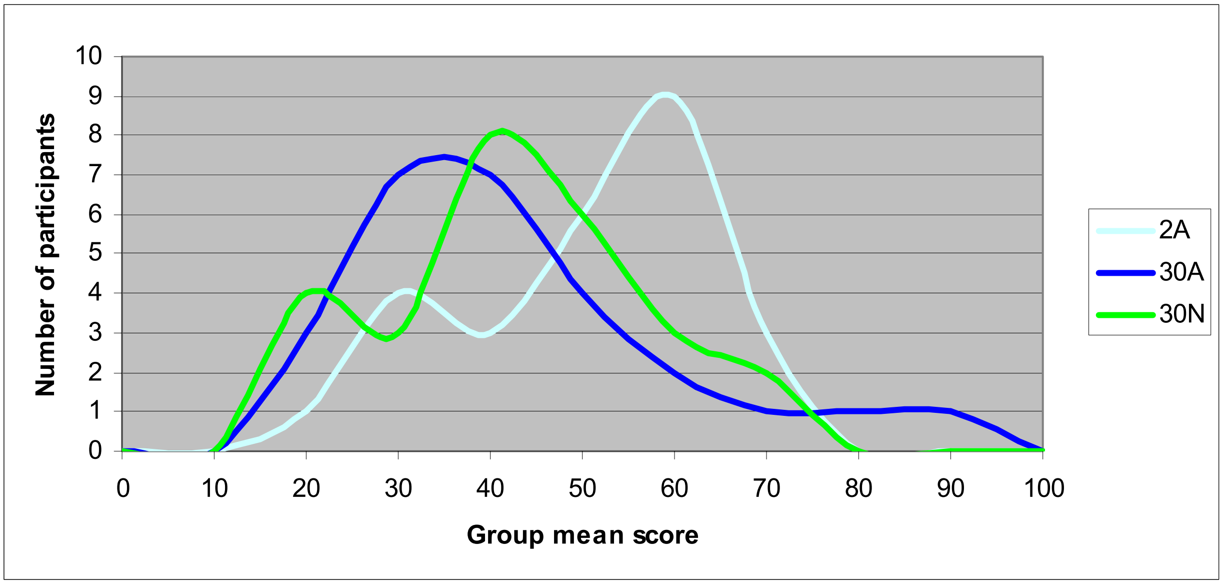


Figure 17: Experiment 2 - Frequency distribution of individual correctness scores (resolution: 10) (N = 26).

Figure 17 shows the frequency distribution of individual correctness scores. The leftward shift of the graph with depth is a clear indication of a change in performance. This shift shows a rather uniform deterioration in the correct scores of participants as related to depth.

3.2.2. Metacognitive performance

Judgment of learning (JOL) for the Word-pair encoding was defined as the mean percent of JOL assessed under water while being exposed to the 50 word-pairs, and was measured after each pair. The group-mean percent of JOL (Table 6) was 48.66 for the shallow water dive, a range of 14.14 to 74.22, 49.49 for the deep dive on air, a range of 22.10 to 83.72, and 51.04 for the nitrox deep dive, a range of 15.66 to 84.08. The one-way ANOVA on diving conditions (2A, 30A, 30N) showed no significant difference between the JOL measurements (F < 1) (see Figure 14).

3.2.3. Metacognitive measurements

As in Experiment 1, the basic confidence scores derived from metacognitive and cognitive performance – metacognitive measurements – were calculated: calibration, and resolution, which included gamma-correlation and discrimination index.

Calibration reflects the overall correspondence between confidence and performance. The group-mean calibration levels, as outlined in Table 6, were 1.51 for shallow water, with a range of -28.8 to 27.26, 10.8 for deep dive with air, with a range of -11.9 to 43.48, and 12.04 for deep dive with nitrox with a range of -25.46 to 44.9. One-way ANOVA between the diving condition (2 msw air, 30 msw nitrox$_{36}$, 30 msw air) yielded a significant effect on the bias score F(2,50) = 8.55, MSE = 100.72, p < 0.001. Post hoc comparisons (Scheffe) showed significant differences between 2 msw and both 30 msw dives in favor of the shallow dives, indicating that at the depth of 30 msw participants showed over-confidence, unlike the 2 msw dives, where participants were well calibrated. No significant difference was found between the 30 msw nitrox and the 30 msw air dives (see Figure 15).

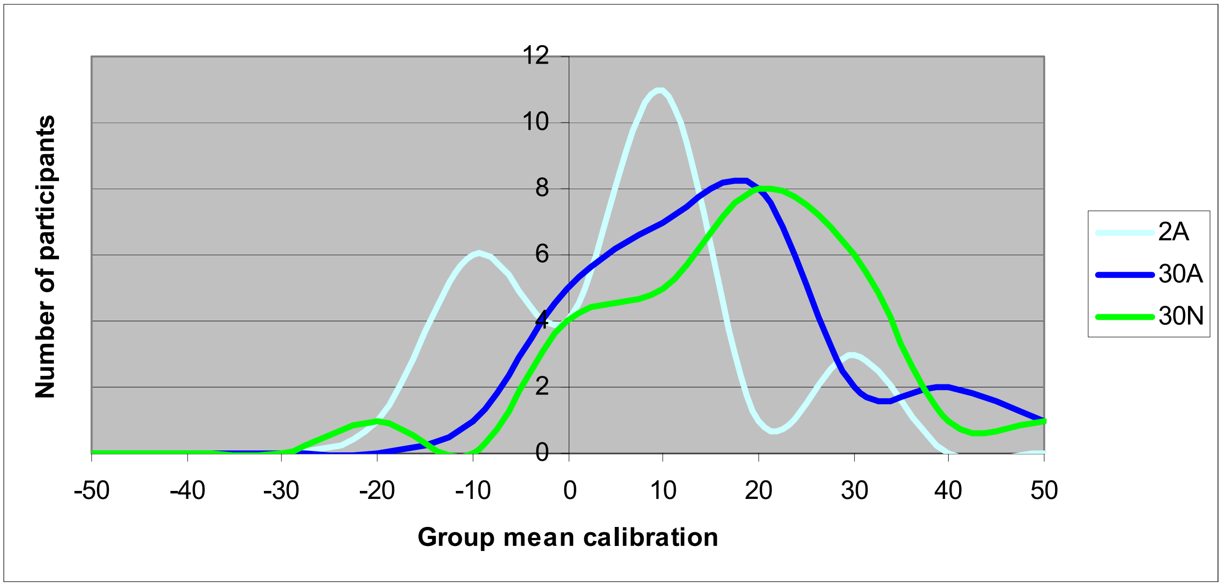


Figure 18: Experiment 2 - Frequency destribution of individual JOL calibration values (resolution: 10) (N = 26).

Figure 18 shows the frequency distribution of individual JOL calibration values. A rightward shift of the two depth graphs is readily apparent. Notably, naturally under-confident participants became calibrated or even over-confident at depth. None of the participants became less confident at depth!

Resolution

Gamma-Correlation: The group-mean gamma-correlation, as can be seen in Table 6, was 0.6 for shallow water, with a range of 0.28 to 0.9, 0.58 for deep dive on air with a range of 0.2 to 0.84, and 0.53 for deep dive on nitrox, with a range of 0.06 to 0.89. The one-way ANOVA on diving conditions (2 msw air, 30 msw nitrox$_{36}$, 30 msw air) shows an insignificant effect for the mean gamma-correlation ($F<1$).

Discrimination: The group-mean discrimination index, shown in Table 6, was 27.79 for shallow water, with a range of 8.37 to 64.29, 25.91 for deep dive on air, with a range of 1.54 to 57.52, and 23.02 for deep dive on nitrox, with a range of 0.79 to 52.62. One-way ANOVA between the diving condition (2 msw air, 30 msw nitrox$_{36}$, 30 msw air) yielded no insignificant effect on the bias score: $F(2,50) = 1.85$, $MSE = 0.023$, $p < 0.17$.

3.2.4. Consistency of cognitive and metacognitive performance over the three diving conditions

The correlation in Experiment 2 between the three diving conditions (2A, 30A and 30N) as regards both performance and judgment, found all combinations to be significantly correlated, some with high coefficients (Table 7). This means that a participant who performed better on 2 msw air also did so in the other conditions, and also that one who was less well calibrated in one deep dive was also less well calibrated in the other.

Table 7: Experiment 2 - Participant correlation coefficients between diving conditions (N = 26).

Correlation	**2A-30A**	**2A-30N**	**30A-30N**
Correct	.66**	.71**	.80**
JOL	.65**	.77**	.69**
Calibration	.43*	.51*	.60**
Discrimination	.50*	.53*	.47*

** $p < 0.001$ * $p < 0.05$

3.2.5. Summary

The Word-pair Test results shown in Table 6, demonstrate a clear difference between 2 msw dives and 30 msw dives. The participants managed to give more correct answers at 2 msw than at 30 msw, though their JOL remained at approximately the same level, causing a drop in the level of calibration in the direction of over-confidence. There was no significant difference between nitrox and air.

Experiment 2 also demonstrates a positive autocorrelation between the individual's ability on all dives and conditions. These correlations are seen both in the cognition and metacognition measurements. Some of the trials were widely separated in time, meaning that there is a consistency over time in a participant's psycho-physiological profile, and in the manner that such a profile is affected by depth.

CHAPTER 4: DISCUSSION

Three experiments were carried onshore and underwater (over 12 minutes): Experiment 1, the Faces Test, tested visual memory and the ability to recall a face to which the participant had been exposed to on land, underwater. Experiment 2, the Word-pair Test, examined literal memory, in which exposure and recall were performed underwater. The Picture Test, which was later disqualified due to technical problems, was designed to examine visual memory to which the participant had been exposed to under water and was asked to recall on land.

Four main issues were examined throughout this research, and will be discussed: the suitability of the cognitive and metacognitive tests for the particular task, the influence of depth, the influence of the gas mixture, and the possibility of devising a future screening tool for divers working in the realm of nitrogen narcosis.

4.1. <u>Depth effects using cognition and metacognition measurements</u>

Davis *et al.* (1972) examined different skills at a depth of 30 msw. They succeeded in demonstrating significant deterioration in manual dexterity, sentence comprehension and arithmetic, but not in memory. In this research only one experiment out of the three showed reliability and significant results. The low results of dive correlation coefficients of the correctness scores of Experiment 1 imply an unreliable cognitive experiment, although metacognition was seemingly consistent. Experiment 3 was disqualified due to logistical problems. This is the reason that this discussion focuses mostly on the results of Experiment 2.

Experiment 2 was the more reliable and sensitive experiment, and proved that depth has a major influence on the memory process by showing a clear cognitive performance decrease as depth increases. However, the group-mean JOL remains the same as in the shallow dive. This means that the participants were not capable of foreseeing their decreased performance. A similar pattern was found in a study of how people judge their knowledge about forgetting by Koriat *et al.* (2004). They asked people to estimate how many word-pairs they would remember immediately, after a day, and after a week, and found that people gave the same estimates for the three time conditions, although their actual correctness score deteriorated with time.

The calibration results of Experiments 1 and 2 are an indication of confidence. Calibration is the difference between the average of assessments and the proportion of correct answers (Lichtenstein *et al.*, 1982; Koriat and Goldsmith, 1996). In the Word-pair Test, the calibration results showed a shift of divers' performance by an average of 10% positive difference, indicating over-confidence. This indicates that this degree of impairment, although quite substantial, does not seem to reach the consciousness of the participants as a group.

Although calibration in Experiment 2 showed significant over-confidence, neither the gamma-correlation nor the discrimination showed a significant decrease, although there was a downward trend. Had the participants erred in the easy questions, for which they gave high JOL scores, we would have expected to see a significant drop in resolution, so apparently the extra errors at depth were in medium hard questions (with scores close to 50%), that influenced resolution only slightly. A greater depth might have made the trend significant.

This finding of over-confidence while under the influence of nitrogen narcosis is well known in qualitative studies (Nevo and Breitstein 1999). In this research, over-confidence is demonstrated quantitatively. This would not have been possible without using metacognition measurements. Nevo and Breitstein (1999) summed up all memory experiments as showing a deterioration of 5-10% at 30 msw. This research, in Experiment 2, achieved a clear 20% deterioration. This high percentage again shows the strength and sensitivity of the Word-pair Test.

Baddeley and Idzikowski (1985), who examined the effect of submersion on divers, summed up their research findings by claiming that at depths up to 30 meters, more than half of the impairment to cognitive functioning while diving stems from water effect. The deterioration is not only between dry land and diving, but also between shallow and deep dives. If this were the case, then one should expect deterioration in the results of the Faces Test as well. The design of the present experiment neutralizes any immersion effects by using a wet control. The results of the Faces Test show no change in cognitive results with depth.

The confidence of the participants in Experiment 1 was the same throughout the entire process and at all three conditions. Confidence was not tested in Experiment 2, so it is not certain whether it too would have shown over-confidence. In addition, in Experiment 1 encoding was done onshore in dry conditions, and never at depth as in Experiment 2. Had this been tested, an effect of depth on this experiment might yet have been found. This factor supports the Baddeley and Idzikowski (1985) research. However, to reach this conclusion the experiment should be repeated using the same time gaps.

Unlike Nelson *et al*. (1990), who found no effect on recall on Mount Everest, this research found a clear deterioration at our 'high altitude', or depth of 30 msw. Nelson acknowledges this by citing Cohen's (1989) conclusion that "memory (recall) in the real world proves remarkably efficient and resilient" (p. 222). Contrary to Nelson, our findings show a clear deterioration in memory abilities in Experiment 2.

All this leads to the conclusion that the results of Experiment 2 show a clear deterioration of cognitive abilities at a depth of 30 msw. The participants were not able to detect this deterioration at depth, and were consistent in their JOL. The difference between the JOL and the correctness is the confidence, which is estimated at +10%. This shows a clear over-confidence at depth. A lesson to be learnt from all this is that moderate depths with subtle narcotic effects, which, as has been shown here, may still cause marked underperformance in some tasks, are more dangerous by not being accompanied by divers' awareness in deeper dives, where narcosis is subjectively felt and taken into account.

4.2. Gas mixture influence

This research examined the different effects of nitrogen narcosis on divers breathing air and nitrox, using cognition and metacognition measurements. There are three main claims regarding nitrox and nitrogen narcosis: the U.S. Navy (Anon, 1999) and a few diving organizations (e.g. Israeli Diving Federation, ANDI), who claim that nitrox reduces the effect of nitrogen narcosis, several researchers who claim that oxygen has the same narcotic effect as the nitrogen in the mixture, based on an unverified hypothesis (see Linnarsson *et al.*, 1990; Joiner, 2001b; Edmonds, 2002; Bennett and Rostein, 2003), and some organizations, (e.g. Divers Alert Network – DAN), which found it difficult to determine, and declared that there was too little research on this issue (Hamilton, 2000).

The results from Experiment 2 reflect a clear similarity in the effects of air and nitrox. Further examination of the experimental setup may lead to some deductions before reaching the conclusion that air and nitrox cause a similar degree of nitrogen narcosis, at least at 30 meters.

The 30-meter depth is the theoretical threshold at which nitrogen narcosis occurs (Nevo and Breitstein 1999; Joiner 2001a; Bennett and Rostein 2003,). The guidelines of the University Diving Officer and the diving certifications of the participants determined 30 meters to be the floor depth for the experiment. It is possible that at a greater depth and degree of narcosis, a mixture effect would have become apparent.

As previously discussed, nitrogen narcosis is individual, and affected by various variables which were not controlled by the researcher: fatigue, alcohol (consumed the night before), mental stress, etc. (Nevo and Breitstein, 1999). These uncontrolled variables may have biased individual dives, and due to the small number of participants, by increasing the variance, may have contributed to the masking of any mixture effect.

Nitrox$_{36}$ differs only slightly from air in the percentage of nitrogen, which may explain the similarity of the findings. It may be that if nitrox$_{40}$ had been used, the difference might have been significant. However, oxygen toxicity had to be considered in relation to the depth.

Finally, it may be acknowledged that nitrox and air indeed cause an identical degree of nitrogen narcosis at this depth. This is confirmed by results of Linnarsson *et al*. (1990), who sum up their hyperbaric experiment concerning nitrogen narcosis by stating "Our results suggest that the degree of narcosis is not significantly ameliorated when a substantial part of the N_2 is substituted with O_2, and are therefore in favor of the notion that O_2 contributes to the narcotic action of hyperbaric air." (Linnarsson *et al*., 1990, p. 341).

There is still also the possibility that the shallow dive was not a proper control, and that the deep dives contained other elements other than narcosis (poorer visibility, less insulation by the wet suit, general apprehension) that caused the deterioration. However, the fact that none of the participants lost confidence at depth seems to speak against apprehension and discomfort as major contributors.

Edmonds (2002) claimed that on theoretical grounds, assuming N_2 pressure was the sole cause of nitrogen narcosis, a Nitrox$_{36}$ dive at a depth of 30 meters would be "narcotically equivalent" to an air dive with at 23 msw, i.e. about nil. This has never been verified successfully (see Linnarsson *et al*., 1990), and is also refuted by the present experiment.

An examination of the results of nitrox and air at depth of 30 msw demonstrates a striking similarity, in both cognitive and metacognitive measurements. At least, at the studied depth, nitrogen and oxygen seem to have an equal narcotic effect. Whether this similarity will hold at greater depths and higher percentage of oxygen in the mixture, awaits further research.

4.3. Screening tool

In Experiment 1, the metacognitive performance was found to be unreliable, using the face recognition task. In contrast, the metacognitive accuracy in terms of calibration, as well as in terms of resolution, was found to be reliable in Experiment 2, using the Word-pairs recall task. The results of Experiment 2 indicate that a participant who tends to show over-confidence on 2 msw air will also show over-confidence in other conditions as well. In addition, a person who is able to discriminate between what he/she knows and what he/she does not on 2 msw air, will succeed in discriminating between those two on other diving conditions to the same extent. These results of consistency in cognitive performance and calibration are consistent with previous findings on general knowledge (Schraw *et al.*, 1995; Thompson and Mason, 1996) and face recognition (Thompson and Mason, 1996). However, the results of consistency of resolution are new, (found before only by Sheffer, 2003), and suggest that the tendency to overestimate one's performance might constitute a stable personal characteristic, and that the ability to distinguish between what one knows and what one does not, is consistent.

Since the performance, and more important, the accuracy of self-assessments, on the surface as well as under water, were found to be stable, personal qualities over time and at different diving conditions, these tasks can serve as an assessment and screening tool for professional deep divers. By using sensitive metacognition tests, in combination with the Word-pair Test, testers can assess the individual's personal performance and judgment at depth, and may weed out participants who demonstrate a combination of high sensitivity to nitrogen narcosis, shown by a marked decrease in the cognitive score at depth, and a decreased awareness and/or true overconfidence.

There is need for extensive further use of the metacognition measurements in deep diving in order to collect a large enough database from which to determine the minimum standards a diving candidate should achieve in order to prove his/her ability in this regard. Since there is also correlation between depths, there is a possibility of constructing a 'dry' test on land, which would reflect the underwater performance.

4.4. Test suitability

While planning this experiment there were many unknown factors. One big question mark was regarding which memory test should be used to reveal nitrogen narcosis at this moderate depth. Another principle task was to try to isolate nitrogen narcosis effects on the encoding and recall phases of the memory test. The former was resolved by a battery of tests, which were used as separate experiments, and the latter by preferentially splitting the test phases on land and under water in the various experiments. The pre-chosen tests were pilot-tried on land, mainly to work out the timing, but there was no way to predict which one would prove to be sufficiently sensitive.

The main problem in constructing an underwater experiment is the tight timetable, which for the depth chosen for this study, gave 20 minutes' bottom time. The 20 minutes had to include the descent of all participants until they were on their knees ready to begin the experiment, a process that sometimes took eight expensive minutes. Hence, the total time for the experiment should not have exceeded 12 minutes. The maximum depth for the experiment was 30 msw. The effect of nitrogen narcosis at this depth ranges from none to mild, and the first thing to influence the diver suffering from nitrogen narcosis is a mild impairment of performance in unpracticed tasks. The tasks were indeed unpracticed, and in order to get an extra dimension to enhance the depth effect as well as the differences between mixtures, a metacognition option was added to each question. The combination of cognitive and metacognitive measurements (the latter used for the first time under water) were expected to reveal the mild impairment in an accurate manner by providing the diver's calibration and resolution, thus quantifying impairment of judgment, be it as sensitive as it may.

Three experiments were carried out during these 12 minutes. Experiment 1, the Faces Test, tested visual memory and the ability to recall underwater a face, which the participant had been exposed to on land. Experiment 2, the Word-pair Test, examined literal memory, in which exposure and recall were performed under water. Experiment 3, the Picture Test, which was later disqualified due to technical problems, was designed to examine visual memory to which the participant had been exposed to underwater and asked to recall on land.

As it turned out, the first test, involving a recognition type recall, proved unaffected by this depth, while the second, free recall, proved to be sensitive enough to show depth effects. Experiment 1 also showed a perfect calibration, again, probably on account of involving a recognition type recall. Another proof for the insensitivity of Experiment 1 may be deduced from Godden and Baddeley's (1975) experiment concerning the effect of submersion on memory, which examined participant's memory at the surface and at 5 meters under water. The results showed that there was no difference between study on land or under water. There was, however, a very clear context-dependency effect. If these results are applied to the present experiments, we should have expected Experiment 1 to do worse and to be less well calibrated, due to the fact that encoding and recall occurred in different environments.

As it is evident that Experiment 2 (the Word-pair Test) was the one best suited for showing mild performance impairment at moderate depths, it is recommended that further studies use this test in all of the above three versions, which would also offer an opportunity to repeat the results of Godden and Baddeley. As Experiment 3 failed technically, its usefulness and sensitivity could not be evaluated, but, being also a free recall test, it still holds promise, and is a viable candidate for future studies

4.5. Lessons

Several problems surfaced while performing this study. There is as much to learn from the mistakes as from the results. Due to the scheduling problems and budgetary and academic problems, the experiment that was meant to be completed in one month (December) continued into January and February. The main problem associated with this time delay was the drop in water temperature, starting at 21 ^{0}C and dropping to 15 ^{0}C later on. The interesting fact was that, when checked, the lower temperature seemed to have made the participants more focused, and actually brought about a better performance (this requires further research). It is important to state that sea conditions (e.g. currents, visibility, etc.) were similar throughout all dives, except for the issue of temperature. The staggered order of dive types largely cancelled the temperature effect.

The offshore dives were from a buoy anchored to the sea bottom, which, due to surface currents, moved around the anchor. The descent after tying up to the buoy was with the aid of a plummet with a rope in order to shorten descent time and make it easier to find the boat upon ascent. Due to the movement of the buoy, bottom depth was not the same for all dives, and ranged from 28 to 30 msw.

The experimental timetable was somewhat flexible, dependent on various factors. Among these were unexpected delays, which kept the boat from leaving according to schedule and arriving at the exact time at the diving site, or problems with equalizing pressure, causing delays in descent. All of this made it impossible to fix the time between the encoding phase on the surface and the start of the recall test underwater (40 to 60 minutes). The underwater experiment was precisely timed, thanks to the Timekeeper.

A few tests had to be cancelled because of version repetition as a result of a mistake made by the Experiment Manager. A participant was not allowed to answer the same test version twice, and if he/she did, the test was cancelled. Another reason to cancel a test was a participant's misunderstanding of the instructions. A cancelled test did not interfere with the other experiments, which were completed successfully. Only completed experiments (Faces or Word-pair) were taken into consideration.

The experiments were constructed to a very tight time schedule and were examined in a pilot study. The pilot helped to improve the test versions so that they would present an equal degree of difficulty. The tests were checked statistically for deviation between versions, and were found equally leveled. However, there were a few mistakes which were not detected in the pilot, such as ambiguous questions and unclear phrasing. The picture Test was cancelled because unclear print made it almost impossible to answer the questions; however, the relevance of visual memory in diving may make it another test of choice, once problems are overcome.

4.6. Future Research

This research has raised some ideas that need to be examined further.

In Experiment 2, participants were unable to foresee their impairment of cognitive abilities at 30 msw, unlike at 2 msw, where the correctness and JOL were almost identical. It would be interesting to find the threshold depth at which one (or a group) starts to become aware of this effect of nitrogen narcosis.

Experiment 2 shows a clear effect of depth. However, the potential separate depth effects on encoding and recall phases were not isolated in this experiment. This is important information that should be gained in future studies by administering the Word-pair and possibly the Picture Test in the dry-wet, wet-wet and wet-dry versions, using the same time gaps between encoding and recall.

In addition, confidence was not assessed in Experiment 2, yet it should be in future designs, as it is possible that it too would not have changed, but also possible that it would have decreased, rendering participants better calibrated than in their JOL. In that case, the whole argument of over-confidence may require reconsideration.

CHAPTER 5: SUMMARY

This research is a preliminary attempt to answer the question of whether nitrox and air differ in their degree of nitrogen narcosis. Being informed of the fact that both performance and judgment of a diver may be equally impaired during a nitrox dive and an air dive at the same depth, may contribute to divers' awareness and to safer diving. The use of metacognition measurements during deep dives has never before been attempted, and it provides a new tool in understanding psycho-physiological aspects of diving. At the start of this research, there was as much unknown as known. At its conclusion, it can positively be said that Experiment 2, the Word-pair Test, is the one of choice for further research. This test showed sensitivity sufficient to demonstrate a nitrogen narcosis effect at the chosen moderate depth, and holds further promise if applied at greater depths. The lack of awareness of participants to their decreased performance at depth was apparent with this test, and further studies at greater depth may demonstrate true over-confidence as one of the effects of narcosis.

Although the purpose of this research was to assess nitrox ability to alleviate the nitrogen narcosis effect, a clear verdict on this issue was not reached. There was enough power in the experiments to discern a difference, had there been one, so one could say with reasonable confidence that at 30 msw nitrox and air have a similar narcotic affect. This finding should be verified at a greater depth and/or with a higher percentage of oxygen.

Future research should try to separate effects of nitrogen narcosis on the encoding and recall of data, a task unfulfilled in the present study. An experiment should be devised along the same lines as the present one, with the three combinations of land-underwater split dives, but using the Word-pair, and possibly a better version of the Picture Test, in all three combinations.

The participants showed consistency over time in their performance and judgment. This means that a participant who performed poorly at depth as well as showing poor judgment, continues to do so on both deep dives. This consistency is a prerequisite for forming a screening tool for professional deep divers, or help spot the divers best suited (least affected performance and best calibrated) for complex deep underwater missions. A hyperbaric chamber should help give better results, since it can allow extension of the depth, the timetable and the percentage of oxygen in the gas mixtures, all key factors in this experiment. Finally, the collection of more data on divers' cognitive and metacognitive behavior at depth may help tackle the overconfidence tendency under water, thus making diving safer.

CHAPTER 6: BIBLIOGRAPHY

Anon. (1999). Nitrogen-Oxygen Diving Operations. In: U.S. Navy Diving Manual, Supervisor of Diving, Volume 2, 10.

Adolfson, J. (1965). Deterioration of mental and motor functions in hyperbaric air. Scandinavian Journal of Psychology, 6, 26-32.

Adolfson, J. and Muren, A. (1965). Air breathing at 13 atmospheres: Psychological and physiological observations. Sartrych für Forsvasmedicin, I, 31-37.

Bachrach, A.J. (1982). A short history of man in the sea. In: Bennett, P.B. and D.H. Elliott (Eds.) Physiology and medicine of diving. 3rd edition. Best Publishing Co. San Pedro, CA. 1-14.

Baddeley, A.D., (1966). Influence of depth on the manual dexterity of free divers. Journal of Applied Psychology, 50, 81-85.

Baddeley, A.D., de Figueredo, J.W., Hawkswell-Curtis, J.W. and Williams, A.N. (1968). Nitrogen narcosis and performance underwater. Ergonomics, 11(2), 157-164.

Baddeley, A.D., and Idzikowski, C. (1985) Anxiety, manual dexterity and diver performance. Ergonomics, 28(10), 1475-1482

Banks, W.W., Berghage, T.E. and Heaney, D.M., (1979). Visual recognition thresholds in a compressed air environment. Aviation, Space and Environmental Medicine, 50, 1003-1006.

Bennett, P.B. (1988). Nitrogen Narcosis. In: Waite, C.L., (Ed.). Case histories of diving and hyperbaric accidents. UHMS, Maryland, 219-224

Bennett, P.B. (1997). Inert gas narcosis and high pressure nervous syndrome. In: Bove, A.A. (Ed.). Bove and Davis' Diving Medicine. 3rd edition. Saunders, Philadelphia, PA, 117-130.

Bennett, P.B., Ackles, K.N. and Cripps, V.J. (1969). Effects of hyperbaric nitrogen and oxygen on auditory evoked responses in man. Aerospace Medicine, 40, 521-525.

Bennett, P.B. and Ackles K.N. (1970). The narcotic effects of hyperbaric oxygen. In: Wada, J. and T. Iwa (eds.). Proceedings of the Fourth International Congress on Hyperbaric Medicine. Igaku Shoin Ltd: Sapporo, Japan. 74-79.

Bennett, P.B. and Rostein, J.C. (2003). Inert gas narcosis. In: Bruback, A.O. and Neuman, T.S. (eds.). Bennett and Elliott's Physiology and Medicine of Diving. 5th Edition. Elsevier Science, Bodmin, Cornwall, 300-322.

Biersner R.J., Hall, D.A., Linaweaver, P.G. and Neuman, T.S. (1978). Diving experience and emotional factors related to the psychomotor effects of nitrogen narcosis. Aviation, Space and Environmental Medicine, 49, 959-962.

Brown, R. and McNeil, D. (1966). The "tip of the tongue" phenomenon. Journal of Verbal Learning and Verbal Behavior, 5, 325-337.

Cohen, G. (1989). Memory in the Real World. London: Lawrence Erlbaum Associates, 247 pages

Davis, F.M., Osborne, J.P., Baddeley, A.D. & Graham, M.F., (1972). Diver performance: Nitrogen narcosis and anxiety. Aerospace Medicine, 43(10) 1079-1082

Doolette, J.D. (2003). Inert gas narcosis. In: T. Mount, (Ed.) Technical Diver Encyclopedia, 2nd edition. IANTD, Miami Shores, FL, 104- 111

Edmonds, C. (2002). Technical diving problems. In: Edmonds, C., Lowry , C., Pennefeather, J. and Walker, R. (Eds.). Diving and Subaquatic Medicine, 4th Edition. Arnold Publications, London, 625-633.

Flavell, J.H. (1976). Metacognitive aspects of problem solving. In: Resnick, L.B, (ed.). The nature of intelligence. Erlbaum Publ. Hillsdale, NJ, 231-235.

Fowler, B., (1973), Effect of hyperbaric air on short-term and long term memory. Aerospace Medicine, 1017-1022

Godden, D. and Baddeley, A.D. (1975). Context-dependent memory in two natural environments: On land and under water. British Journal of Psychology, 66, 325-331

Glenberg, A.M., and Epstein, W. (1987). Inexpert calibration of comprehension. Memory and Cognition, 15(1), 84-93

Hart, J.T. (1965). Memory and the feeling of knowing experience. Journal of Educational Psychology, 56(4), 208-216.

Hamilton, R.W. (1997). Mixed-gas diving. In: Bove, A..A. (ed.). Bove and Davis' Diving Medicine. 3rd edition. Saunders, Philadelphia, PA, 38-64.

Hamilton, R.W. (2000). A brief history of diving with oxygen-enriched air using open circuit apparatus. In: Lang, M.A. (ed.). DAN nitrox workshop proceedings. Divers Alert Network. Durham, NC, 5-18.

Hornbein T. (1965). Everest: the west ridge. The Mountaineers Books, Seattle.

Joiner, J.T. (2001). Diving physiology. In: Joiner, J.T. (ed.). NOAA Diving Manual, 4th Edition. Best Publishing Co., Flagstaff, AZ, 3.20-3.21

Joiner, J.T. (2001). Nitrox diving. In: Joiner, J.T. (ed.). NOAA Diving Manual, 4th Edition. Best Publishing Co., Flagstaff, AZ, 15.1-15.28

Keren, G. (1988). On the ability of monitoring non-veridical perceptions and uncertain knowledge: some calibration studies. Acta Psychologica, 67(2), 95-119.

Kiesling, R.J., and Maag, C.H., (1962). performance impairment as a function of nitrogen narcosis. Journal of Applied Psychology, 46(2), 91-95.

Kindwall, E.P. (1995). Decompression sickness. In: Elliott, D.H. (Ed.) Hyperbaric medicine practice. Best Publishing Company, San Pedro, CA, 312-326.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: further evidence for the accessibility model. Journal of Experimental Psychology: General, 124(3), 311-333.

Koriat, A. (1997). Monitoring one's own knowledge during exposure: A cue-utilization approach to judgments of learning. Journal of Experimental Psychology: General, 126(4), 349-370.

Koriat, A., and Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. Psychological Reviews, 103(3), 490-517.

Koriat, A., and Goldsmith, M. (1998). The role of metacognitive processes in the regulation of memory performance. In: Mazzoni, G. and T.O. Nelson. (eds.). Metacognition and cognitive neuropsychology: Monitoring and control processes. Lawrence Erlbaum, Mahwah, NJ, London, 97-118

Koriat, A., Sheffer, L. and Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgment of learning exhibits increased underconfidence with practice. Journal of Experimental Psychology: General, 131(2), 147-162.

Lichtenstein, S., Fischoff, B. and Phillips, L.D. (1982). Calibration of probabilities: the state of the art to 1980. In: Kahneman, D., P. Slovic and A. Tversky (Eds.). Judgment under uncertainty: Heuristics and biases. Cambridge University Press, Cambridge, England, 306-334

Linnarsson, D., Ostlund, A., Sporrong, A., Lind, F. and Hamilton R.W. (1990) Does oxygen contribute to the narcotic action of hyperbaric air? In: W. Sterk and L. Geeraedts, eds. Proceedings XVIth Meeting of the European Undersea Biomedical Society, Amsterdam: Foundation for Hyperbaric Medicine, 339-341.

Lowry, C., (2002). Inert gas narcosis In: Edmonds, C. Lowry , J. Pennefeather and R.Walker (Eds). Diving and Subaquatic Medicine, 4th edition. Arnold Publications, London, 183-193.

Mears, J.D., and Clearly, P.J., (1980). Anxiety as a factor in underwater performance. Ergonomics, 23(6), 549-557.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. Journal of Verbal Learning and Verbal Behavior, 1, 1-21.

Monteiro, M.G., Hernandez, W., Figlie, N.B., Takahashi, E. and Korukian, M. (1996). Comparison between participative feelings to alcohol and nitrogen narcosis: A pilot exposure. Alcohol. 13(1)m, 75-78.

Nelson, T. O.(1984). A comparison of current measures of the accuracy of feeling of knowing predictions. Psychological Bulletin, 95(1), 109-133.

Nelson, T. O. (1988). Predictive accuracy of the feeling of knowing across different criterion tasks and across different participant populations and individuals. In: Morris, P.E., M.M. Grunberg, and R.N. Sykes (eds.). Practical aspects of memory: Current research and issues. Vol. 1, Wiley, NY, 190-196.

Nelson, T. O., Dunlosky, J., White, M. D., Steinberg, J., Townes, D. B. and Anderson, D. (1990). Cognition and metacognition at extreme altitudes on Mount Everest. Journal of Experimental Psychology 119, 367-374.

Nevo, B. and Breitstein, S. (1999). Nitrogen Narcosis. In: B. Nevo and S. Breitstein. Psychological and Behavioral Aspects of Diving: Best Publishing Company, AZ, 49-55

Philip, R.B., Fields, G.N. and Roberts, W.A., (1989). Memory deficit caused by compressed air equivalent to 36 meters of seawater. Journal of Applied Psychology, 74(3), 443-446.

Phillips, C.J., (1984). Cognitive performance in sports SCUBA. Perceptual and Motor Skills, 59, 635-646.

Pulak C. (1998). The Uluburun shipwreck: An overview. The International Journal of Nautical Archaeology. 27, 188-224.

Rubinsten, O., Anaki, D., Henik, A., Drori. S. and Faran, Y., (2005). Norms for free associations in the Hebrew language. In A. Henik, O. Rubinsten, and D. Anaki (Eds.), Word norms for the Hebrew language (in Hebrew) Ben Gurion University of the Negev.

Rutkowski D. (1996). Nitrox Manual. International association of nitrox and technical divers. Key Largo, Florida.

Schraw, G., Dunkle, M.E., Bendixen, L.D. and Roedel, T.D. (1995) Does a general monitoring skill exist? Journal of Educational Psychology, 87(3), 434-444.

Sheffer L., (2003). The reliability and structure of metacognitive skills and their relationship to cognitive performance. Ph.D. Thesis. University of Haifa, (in Hebrew).

Thompson, W.B., and Mason, S.E. (1996). Instability of individual differences in the association between confidence judgment and memory performance. Memory and Cognition, 24(2), 226-234.

Tulving, E. (1974). Cue dependent forgetting. American Scientist, 62, 74-82

Vernes, R.J., and Darragh, A., (1982). Endocrine reactions and cognitive performance at 60 meters hyperbaric pressure. Correlations with perceptual defense reactions. Scandinavian Journal of Psychology, 23, 193-199.

Wilmshurst, P. (1998). ABC of oxygen and oxygen diving. British Medical Journal. 317, 996-999.

Yaniv, I., Yates, J.F., and Smith, J.K. (1991). Measures of discrimination skill in probabilistic judgment. Psychological Bulletin, 110(3), 611-617

.