Numerons and Vision IAS

Psychology - Paper I IAS (Mains), 2012

Reference Material Support

Dear Aspirants,

Psychology as an optional for IAS Mains has come a long way since it started gaining popularity in the latter half of 90s. Claiming to be one of the most popular optionals, students from diverse backgrounds have fared better and better with Psychology in Civil Services Examination, mostly securing further rank improvements due course of time. A major turnaround came in 2008, when focus shifted on conceptual clarity with simplicity of expression. Year of 2009 saw the toughest questions set ever, while 2010 and 2011 framed a mixed bag of questions in both the papers. In an era when being unpredictable is a way to flex muscles, munching over the next move of UPSC might not prove a much parsimonious debate. A sense catching over the needs of time is to cover the entire syllabus, keeping in view the recent trends and latest developments relating to every topic of importance. Conceptual clarity has always been essential for a rank defining score so, not advisable to compromise on this front. Writing an answer with a fresh appeal and dynamic framework will keep one a step ahead.

This year, once again UPSC has been more inclined towards having an experimental approach in Psychology optional, especially Paper-I. Approximately, one third of the marks were allotted to statistical portion in the first paper. Much to the bewilderment of an average student, two-three questions were completely out of sync with the previous trend, most glaring examples should include Neisser's experiment and McCrary-Hunter 'invariance hypothesis'. One might have hard time reaching a logical conclusion behind this unconventional question framing strategy. Second paper was a sigh of relief but this could also be a 'contrast effect'. Being unpredictable might impression a fear of omnipotent, but here UPSC, better be called, setting an example of being directionless. Either of their strategy analysts might have been through manic episode upon being intimitated of institutional preference for a tough paper. Manic episodes could have potentially disastrous outcome, if coupled with self-indulgent pseudo intellectualism. Now a days, explanations are being sought for 'elegant' theories! Seems that some Shakespeare shaking spears in audio-visual merchandising of Psychology paper!!

Advise for students would be to get prepared for such rude shocks. This exam has potential to turn even a hardcore atheist start believing third forces! Not for it being mother of all examinations, but a bad mother!! Perceived reality of toughness here, could well be attributed to non-directional **stochastic processes** of selection and evaluation. Best strategy would still be to focus upon conceptual clarity and simplicity of expression. In response to Psychology paper this year, a detailed reference support material has been prepared that could get one thorough conceptual understanding of the questions framed. This is the first set of two drafts, covering Paper-I. Some of the questions have been referred to various scholarly articles and relevant studies, which are linked with downloadable files in PDF format. Those browsing it online need to just click over the links provided. For those with printed version can visit **Numerons** or **Vision IAS** website for this draft in PDF format with all the activated links. Hope, you would enjoy reading this.

Thanks and best wishes!

Amit Shekhar

www.numerons.com
www.visionias.wordpress.com

2

Section - A

1. Answer the following, each in not more than 150 words: 12x5=60

(a) Elucidate the contribution of Sir Francis Galton in shaping psychology.

TOPIC: *Introduction*
SUBTOPIC: *Historical Antecedents of Psychology*
LEVEL: *Easy*
NATURE: *Fundamental*

VISION IAS

REFERENCE:

Galton had a prolific intellect, and produced over 340 papers and books. He was the first to apply statistical methods to the study of human differences and inheritance of intelligence, and introduced the use of questionnaires and surveys for collecting data on human communities, which he needed for genealogical and biographical works and for his anthropometric studies.

He was a pioneer in eugenics, coining the term itself and the phrase "nature versus nurture". His book Hereditary Genius (1869) was the first social scientific attempt to study genius and greatness.

As an investigator of the human mind, he founded psychometrics (the science of measuring mental faculties) and differential psychology and the lexical hypothesis of personality. He devised a method for classifying fingerprints that proved useful in forensic science. He also conducted research on the power of prayer, concluding it had none by its null effects on the longevity of those prayed for.

Galton was interested at first in the question of whether human ability was hereditary, and proposed to count the number of the relatives of various degrees of eminent men. If the qualities were hereditary, he reasoned, there should be more eminent men among the relatives than among the general population. To test this, he invented the methods of historiometry. Galton obtained extensive data from a broad range of biographical sources which he tabulated and compared in various ways. This pioneering work was described in detail in his book Hereditary Genius in 1869. Here he showed, among other things, that the numbers of eminent relatives dropped off when going from the first degree to the second degree relatives, and from the second degree to the third. He took this as evidence of the inheritance of abilities.

Galton invented the term eugenics in 1883 and set down many of his observations and conclusions in a book, Inquiries into Human Faculty and Its Development. He believed that a scheme of 'marks' for family merit should be defined, and early marriage between families of high rank be encouraged by provision of monetary incentives. He pointed out some of the tendencies in British society, such as the late marriages of eminent people, and the paucity of their children, which he thought were dysgenic. He advocated encouraging eugenic marriages by supplying able couples with incentives to have children. On October 29, 1901, Galton chose to address eugenic issues when he delivered the second Huxley lecture at the Royal Anthropological Institute.

Innovations in statistics and psychological theory

Historiometry

The method used in Hereditary Genius has been described as the first example of historiometry. To bolster these results, and to attempt to make a distinction between 'nature' and 'nurture' (he was the first to apply this phrase to the topic), he devised a questionnaire that he sent out to 190 Fellows of the Royal Society. He tabulated characteristics of their families, such as birth order and the occupation and race of their parents. He attempted to discover whether their interest in science was 'innate' or due to the encouragements of others. The studies were published as a book, English men of science: their nature and nurture, in 1874. In

the end, it promoted the nature versus nurture question, though it did not settle it, and provided some fascinating data on the sociology of scientists of the time.

The Lexical Hypothesis

Sir Francis was the first scientist to recognize what is now known as the Lexical Hypothesis. This is the idea that the most salient and socially relevant personality differences in people's lives will eventually become encoded into language. The hypothesis further suggests that by sampling language, it is possible to derive a comprehensive taxonomy of human personality traits.

The questionnaire

Galton's inquiries into the mind involved detailed recording of people's subjective accounts of whether and how their minds dealt with phenomena such as mental imagery. In order to better elicit this information, he pioneered the use of the questionnaire. In one study, he asked his fellow members of the Royal Society of London to describe mental images that they experienced. In another, he collected in-depth surveys from eminent scientists for a work examining the effects of nature and nurture on the propensity toward scientific thinking.

Theories of perception

Galton went beyond measurement and summary to attempt to explain the phenomena he observed. Among such developments, he proposed an early theory of ranges of sound and hearing, and collected large quantities of anthropometric data from the public through his popular and long-running Anthropometric Laboratory, which he established in 1884 where he studied over 9,000 people. It was not until 1985 that these data were analyzed in their entirety.

Differential psychology

Galton's study of human abilities ultimately led to the foundation of differential psychology and the formulation of the first mental tests.

(b) Discuss the criteria of question-writing in a survey research.

TOPIC: *Research Methods*
SUBTOPIC: *Methods of data collection (interview, observation, questionnaire)*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

Constructing a Questionnaire www.numerons.com

- Constructing a questionnaire involves deciding what information should be sought and how to administer the questionnaire, writing a draft of the questionnaire, pretesting the questionnaire, and concluding with specifying the procedures for its use.
- The wording of questionnaires should be clear and specific using simple, direct, and familiar vocabulary.
- The order in which questions are asked on a questionnaire needs to be considered seriously because the order can affect respondents' answers.

Steps in Preparing a Questionnaire Constructing a questionnaire that will yield reliable and valid measurements is a challenging task. In this section we suggest a series of steps that can help you meet this challenge, especially if you are constructing a questionnaire for the first time as part of a research project.

1. Decide what information should be sought.

4

©VISION IAS
www.visionias.wordpress.com
Contact: +91-9650617807, +91-9968029039

©Numerons
www.numerons.com
Contact: +91-9560344245

2. Decide how to administer the questionnaire.
3. Write a first draft of the questionnaire.
4. Reexamine and revise the questionnaire.
5. Pretest the questionnaire.
6. Edit the questionnaire and specify the procedures for its use.

Step 1. The warning "Watch out for that first step!" is appropriate here. The first step in questionnaire construction—deciding what information is to be sought—should actually be the first step in planning the survey as a whole. This decision, of course, determines the nature of the questions to be included in the questionnaire. It is important to predict the likely results of a proposed questionnaire and decide whether these "findings" would answer the questions of the study. Surveys are frequently done under considerable time pressure, and inexperienced researchers are especially prone to impatience. A poorly conceived questionnaire, however, takes as much time and effort to administer and analyze as does a well-conceived questionnaire. The difference is that a well-constructed questionnaire leads to interpretable results. The best that can be said for a poorly designed one is that it is a good way to learn the importance of careful deliberation in the planning stages.

Step 2. The next step is to decide how to administer the questionnaire. For example, will it be self-administered, or will trained interviewers be using it? This decision is determined primarily by the survey method that has been selected. For instance, if a telephone survey is to be done, trained interviewers will be needed. In designing the questionnaire, one should also consider using items that have been prepared by other researchers. For example, there is no reason to develop your own instrument to assess racial prejudice if a reliable and valid one is already available. Besides, if you use items from a questionnaire that has already been used, you can compare your results directly with those of earlier studies.

Step 3. If you decide that no available instrument suits your needs, you will have to take the third step and write a first draft of your own questionnaire. Guidelines concerning the wording and ordering of questions are presented later in this section.

Step 4. The fourth step in questionnaire construction—reexamining and rewriting—is an essential one. Questions that appear objective and unambiguous to you may strike others as slanted or ambiguous. It is most helpful to have your questionnaire reviewed by experts, both those who have knowledge of survey research methods and those with expertise in the area on which your study is focused. For example, if you are doing a survey of students' attitudes toward the campus food service, it would be advisable to have your questionnaire reviewed by the campus foodservice director. When you are dealing with a controversial topic, it is especially important to have representatives of both sides of the issue screen your questions for possible bias.

Step 5. By far the most critical step in the development of an effective questionnaire is to do a pretest. A pretest involves actually administering the questionnaire to a small sample of respondents under conditions similar to those anticipated in the final administration of the survey. Pretest respondents must also be typical of those to be included in the final sample; it makes little sense to pretest a survey of nursing home residents by administering the questionnaire to college students. There is one way, however, in which a pretest does differ from the final administration of the survey. Respondents should be interviewed at length regarding their reactions to individual questions and to the questionnaire as a whole. This provides information about potentially ambiguous or offensive items.

The pretest should also serve as a "dress rehearsal" for interviewers, who should be closely supervised during this stage to ensure that they understand and adhere to the proper procedures for administering the questionnaire. If major changes have to be made as a result of problems discovered during the pretest, a second pretest may be needed to determine whether these changes solved the problems.

Step 6. After pretesting is completed, the final step is to edit the questionnaire and to specify the procedures to be followed in its final administration. To reach this final step successfully, it is important to consider guidelines for the effective wording of questions and for the ordering of questions.

(c) Bring out the difference between 'sampling error' and 'error in sampling'. How 'sampling error' is reduced?

TOPIC: *Research Methods*
SUBTOPIC: *Major steps in Psychological research*
LEVEL: *Medium*
NATURE: *Fundamental*

REFERENCE:

In statistics, sampling error is the error caused by observing a sample instead of the whole population. The sampling error can be found by subtracting the value of a parameter from the value of a statistic. In nursing research, a sampling error is the difference between a sample statistic used to estimate a population parameter and the actual but unknown value of the parameter (Bunns & Grove, 2009). An estimate of a quantity of interest, such as an average or percentage, will generally be subject to sample-to-sample variation. These variations in the possible sample values of a statistic can theoretically be expressed as sampling errors, although in practice the exact sampling error is typically unknown. Sampling error also refers more broadly to this phenomenon of random sampling variation.

Random sampling, and its derived terms such as sampling error, imply specific procedures for gathering and analyzing data that are rigorously applied as a method for arriving at results considered representative of a given population as a whole. Despite a common misunderstanding, "random" does not mean the same thing as "chance" as this idea is often used in describing situations of uncertainty, nor is it the same as projections based on an assessed probability or frequency. Sampling always refers to a procedure of gathering data from a small aggregation of individuals that is purportedly representative of a larger grouping which must in principle be capable of being measured as a totality. Random sampling is used precisely to ensure a truly representative population from which to draw conclusions, in which the same results would be arrived at if one had included the entirety of the population instead. Random sampling (and sampling error) can only be used to gather information about a single defined point in time. If additional data is gathered (other things remaining constant) then comparison across time periods may be possible. However, this comparison is distinct from any sampling itself. As a method for gathering data within the field of statistics, random sampling is recognized as clearly distinct from the causal process that one is trying to measure. The conducting of research itself may lead to certain outcomes affecting the researched group, but this effect is not what is called sampling error. Sampling error always refers to the recognized limitations of any supposedly representative sample population in reflecting the larger totality, and the error refers only to the discrepancy that may result from judging the whole on the basis of a much smaller number. This is only an "error" in the sense that it would automatically be corrected if the totality were itself assessed. The term has no real meaning outside of statistics.

According to a differing view, a potential example of a sampling error in evolution is genetic drift; a change is a population's allele frequencies due to chance. For example the bottleneck effect; when natural disasters dramatically reduce the size of a population resulting in a small population that may or may not fairly represent the original population. What may make the bottleneck effect a sampling error is that certain alleles, due to natural disaster, are more common while others may disappear completely, making it a potential sampling error. Another example of genetic drift that is a potential sampling error is the founder effect. The founder effect is when a few individuals from a larger population settle a new isolated area. In this instance, there are only a few individuals with little gene variety, making it a potential sampling error.

The likely size of the sampling error can generally be controlled by taking a large enough random sample from the population, although the cost of doing this may be prohibitive; see sample size and statistical power for more detail. If the observations are collected from a random sample, statistical theory provides probabilistic estimates of the likely size of the sampling error for a particular statistic or estimator. These are often expressed in terms of its standard error.

Reducing the Sampling Errors:

1. **By increasing the size of the sample.** The sampling error can be reduced by increasing the sample size. If the sample size n is equal to the population size N, then the sampling error is zero.

2. **By Stratification.** When the population contains homogeneous units, a simple random sample is likely to be representative of the population. But if the population contains dissimilar units, a simple random sample may fail to be representative of all kinds of units, in the population. To improve the result of the sample, the sample design is modified. The population is divided into different groups containing similar units. These groups are called strata. From each group (stratum), a sub-sample is selected in a random manner. Thus all the groups are represented in the sample and sampling error is reduced. It is called stratified-random sampling. The size of the sub-sample from each stratum is frequently in proportion to the size of the stratum. Suppose a population consists of 1000 students out of which 600 are intelligent and 400 are non-intelligent. We are assuming here that we do have this much information about the population. A stratified sample of size $n =$ 100 is to be selected. The size of the stratum is denoted by N_1 and N_2 respectively and the size of the samples from each stratum may be denoted by n_1 and n_2. It is written as under:

Stratum No.	Size of stratum	Size of sample from each stratum
1	$N_1 = 600$	$n_1 = \dfrac{n \times N_1}{N} = \dfrac{100 \times 600}{1000} = 60$
2	$N_2 = 400$	$n_2 = \dfrac{n \times N_2}{N} = \dfrac{100 \times 400}{1000} = 40$
	$N_1 + N_2 = N = 1000$	$n_1 + n_2 = n = 100$

The size of the sample from each stratum has been calculated according to the size of the stratum. This is called proportional allocation. In the above sample design, the sampling fraction in the population is $\dfrac{n}{N} = \dfrac{100}{1000} = \dfrac{1}{10}$ and the sampling fraction in both the strata is also $\dfrac{1}{10}$. Thus this design is also called fixed sampling fraction. This modified sample &sign is frequently used in sample surveys. But this design requires some prior information about the units of the population. On the basis of this information, the population is divided into different strata. If the prior information is not available then the stratification is not applicable.

(d) Describe the role of epigenetic model in determining human behavior.

TOPIC: **Development of Human Behavior**
SUBTOPIC: **Role of genetic and environmental factors in determining human behavior**
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

Epigenetics in psychology helps to explain how nurture shapes nature, where nature refers to biological heredity and nurture refers to virtually everything that occurs during the life-span (e.g., social-experience, diet and nutrition, and exposure to toxins). Epigenetics in psychology provides a framework for understanding how the expression of genes is influenced by experiences and the environment to produce individual differences in behavior, cognition personality, and mental health.

British biologist Conrad Waddington is credited with developing the term epigenetics in 1947. Epi is a Greek term meaning upon or above. Hence, epigenetics reflects the effects that take place upon, above or in addition to genetics. This original definition implied that something aside from genes was involved in defining the phenotype. When Waddington created the term epigenetics, little was known about the expression of genes. Waddington and others eventually realized that development involved networks of interactions between genes. By the 1990s, Hall wrote that epigenetics involves genetic and non-genetic factors that affect gene expression. Hall further noted that the code of heredity includes an extra layer in addition to DNA. By the 2000s, epigenetics was used to describe how experiences shape heritable genetic expression. Psychosocial and environmental factors have been shown to alter these epigenetic mechanisms which, in turn, influence normal and abnormal psychology. The current definition still maintains its continuity to Waddington's, as epigenetics is used to explain how genes and the environment work in tandem. As such, epigenetic changes do not alter or mutate the DNA. Instead, epigenetic mechanisms affect DNA by regulating the expression of genes. Therefore, due to their regulatory function, epigenetic mechanisms are still considered to occur upon or above the DNA, without actually altering it.

Epigenetics across the life-span

Epigenetic changes occur not only in the developing fetus, but also in individuals throughout the human life-span. Epigenetic changes occur cumulatively over the life-span as evidenced by: (i) increased levels of epigenetic changes in older monozygotic (identical) twins compared to younger twins; and, (ii) identical twins who become more epigenetically dissimilar as they age.

Inheritance

Epigenetic modifications can also be inherited from one generation to the next. Specifically, the effects of epigenetics may be transmitted from parents to children through meiosis and mitosis - processes that pass genetic information from the father into the sperm and from the mother into the egg. Hence, subsequent generations may be affected by the epigenetic changes that took place in the parents. Epigenetic modifications in offspring, resulting from parental care-giving environments, may also be passed down from generation to generation. Epigenetic research counters the long-standing notion that children are born with an uncontaminated genetic slate.

Comparison with evolution

According to Darwin's classic theory of evolution, several generations are required for adaptation to occur. Also known as natural selection, this process contributes to humans' (and other species') ability to survive and reproduce in their particular environments. Adaptation may occur through physiological, structural (anatomical) and/or behavioral changes. In the context of natural selection and psychology, more successful (i.e., better adapted) traits and behaviors are more likely to be passed down from an individual to subsequent generations. Thus, most evolutionary psychologists agree that modern human traits and behaviors can be characterized as beneficial adaptations to the environment. However, other theorists argue that some traits are maladaptive. Not to be confused with evolution, epigenetics provides a mechanism for immediate adaptations to ever-changing environments throughout the human life-span.

Individual diversity, stemming from epigenetic differentiation, cannot be attributed to random mutations. Therefore, one view suggests that epigenetics is inconsistent with Darwinian theory of evolution's concept of random mutations resulting in diversity. Although Lamarckism is not appropriate to explain evolution, it might offer insight into epigenetic-related individuation. For example, evolutionary developmental biology accounts for epigenetic and pseudo-Lamarckian mechanisms whereby environmentally induced variation is inherited. Yet another view maintains that epigenetics can be accommodated within a neo-Darwinian framework. In this context, heritable epigenetic differences are viewed as part of long-term development, whereby the number of generations through which epigenetic modifications traverse is dependent on the duration of the influence on the genes.

(e) Discuss the 'pay-off matrix' in relation to the signal detection theory.

TOPIC: *Sensation, Attention and Perception*
SUBTOPIC: *Signal-Detection and Vigilance*
LEVEL: *Difficult*
NATURE: *Fundamental - Applied*

REFERENCE:

An introduction to the concept of 'Pay-off Matrix'

In game theory, **Pay-off Matrix** is a description of a *game*. Unlike extensive form, Pay-off Matrix representations are not graphical *per se*, but rather represent the game by way of a matrix. While this approach can be of greater use in identifying strictly dominated strategies and Nash equilibria, some information is lost as compared to extensive-form representations. The Pay-off Matrix representation of a game includes all perceptible and conceivable strategies, and their corresponding payoffs, of each player.

In static games of complete, perfect information, a Pay-off Matrix representation of a game is a specification of players' strategy spaces and payoff functions. A strategy space for a player is the set of all strategies available to that player, where a strategy is a complete plan of action for every stage of the game, regardless of whether that stage actually arises in play. A payoff function for a player is a mapping from the cross-product of players' strategy spaces to that player's set of payoffs (normally the set of real numbers, where the number represents a cardinal or ordinal utility—often cardinal in the Pay-off Matrix representation) of a player, i.e. the payoff function of a player takes as its input a strategy profile (that is a specification of strategies for every player) and yields a representation of payoff as its output.

An example

The matrix to the right is a Pay-off Matrix representation of a game in which players move simultaneously (or at least do not observe the other player's move before making their own) and receive the payoffs as specified for the combinations of actions played. For example, if player 1 plays top and player 2 plays left, player 1 receives 4 and player 2 receives 3. In each cell, the first number represents the payoff to the row player (in this case player 1), and the second number represents the payoff to the column player (in this case player 2).

A Pay-off Matrix game		
Player 1 \ Player 2	**Player 2 chooses left**	**Player 2 chooses right**
Player 1 chooses top	4, 3	−1, −1
Player 1 chooses bottom	0, 0	3, 4

Other representations

Often symmetric games (where the payoffs do not depend on which player chooses each action) are represented with only one payoff. This is the payoff for the row player. For example, the payoff matrices on the right and left below represent the same game.

Both players		
	Stag	**Hare**

Just row		
	Stag	**Hare**

Stag	3, 3	0, 2		**Stag**	3	0
Hare	2, 0	2, 2		**Hare**	2	2

Uses of Pay-off Matrix

Dominated strategies

The payoff matrix facilitates elimination of dominated strategies, and it is usually used to illustrate this concept. For example, in the prisoner's dilemma (to the right), we can see that each prisoner can either "cooperate" or "defect". If exactly one prisoner defects, he gets off easily and the other prisoner is locked up for good. However, if they both defect, they will both be locked up for longer. One can determine that *Cooperate* is strictly dominated by *Defect*. One must compare the first numbers in each column, in this case 0 > −1 and −2 > −5. This shows that no matter what the column player chooses, the row player does better by choosing *Defect*. Similarly, one compares the second payoff in each row; again 0 > −1 and −2 > −5. This shows that no matter what row does, column does better by choosing *Defect*. This demonstrates the unique Nash equilibrium of this game is (*Defect*, *Defect*).

The Prisoner's Dilemma		
	Cooperate	**Defect**
Cooperate	−1, −1	−5, 0
Defect	0, −5	−2, −2

The 'pay-off matrix' in relation to the signal detection theory (SDT)

PDF File Link - Getting Technical About Awareness

2. Answer the following, each in not more than 250 words: **20x3=60**

(a) Discuss the three basic conditions for using 't'-test of significance. Describe at least five different uses of 't'-test with examples.

TOPIC: *Research Methods*
SUBTOPIC: *Application of statistical technique*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

The Three Basic Conditions for Using 't'-test of Significance

$$T = \frac{Z}{s}$$

Most *t*-test statistics have the form , where Z and s are functions of the data. Typically, Z is designed to be sensitive to the alternative hypothesis (i.e. its magnitude tends to be larger when the alternative hypothesis is true), whereas s is a scaling parameter that allows the distribution of T to be determined.

As an example, in the one-sample t-test $Z = \dfrac{\bar{X}}{\frac{\sigma}{\sqrt{n}}}$, where $\bar{X}$ is the sample mean of the data, n is the sample size, and σ is the population standard deviation of the data; s in the one-sample t-test is $\hat{\sigma}/\sqrt{n}$, where $\hat{\sigma}$ is the sample standard deviation.

The assumptions underlying a t-test are that (Three basic conditions)

1. Z follows a standard normal distribution under the null hypothesis
2. s^2 follows a χ^2 distribution with p degrees of freedom under the null hypothesis, where p is a positive constant
3. Z and s are independent.

In a specific type of t-test, these conditions are consequences of the population being studied, and of the way in which the data are sampled. For example, in the t-test comparing the means of two independent samples, the following assumptions should be met:

- Each of the two populations being compared should follow a normal distribution. This can be tested using a normality test, such as the Shapiro-Wilk or Kolmogorov–Smirnov test, or it can be assessed graphically using a normal quantile plot.
- If using Student's original definition of the t-test, the two populations being compared should have the same variance (testable using F test, Levene's test, Bartlett's test, or the Brown–Forsythe test; or assessable graphically using a Q-Q plot). If the sample sizes in the two groups being compared are equal, Student's original t-test is highly robust to the presence of unequal variances. Welch's t-test is insensitive to equality of the variances regardless of whether the sample sizes are similar.
- The data used to carry out the test should be sampled independently from the two populations being compared. This is in general not testable from the data, but if the data are known to be dependently sampled (i.e. if they were sampled in clusters), then the classical t-tests discussed here may give misleading results.

Uses of t-Test:

Among the most frequently used t-tests are:

- A one-sample location test of whether the mean of a normally distributed population has a value specified in a null hypothesis.
- A two sample location test of the null hypothesis that the means of two normally distributed populations are equal. All such tests are usually called **Student's t-tests**, though strictly speaking that name should only be used if the variances of the two populations are also assumed to be equal; the form of the test used when this assumption is dropped is sometimes called Welch's t-test. These tests are often referred to as "unpaired" or "independent samples" t-tests, as they are typically applied when the statistical units underlying the two samples being compared are non-overlapping.
- A test of the null hypothesis that the difference between two responses measured on the same statistical unit has a mean value of zero. For example, suppose we measure the size of a cancer patient's tumor before and after a treatment. If the treatment is effective, we expect the tumor size for many of the patients to be smaller following the treatment. This is often referred to as the "paired" or "repeated measures" t-test: see paired difference test.
- A test of whether the slope of a regression line differs significantly from 0.

For details see the reference file attached:

PDF File Link: The t-test and Basic Inference Principles

(b) Compare Sperling's experiments with that of Neisser's experiments in the study of sensory memory. What did these two experiments prove?

TOPIC: *Memory*
SUBTOPIC: *Sensory memory, Iconic memory, Echoic memory*
LEVEL: *Difficult*
NATURE: *Fundamental*

REFERENCE:

The discovery of iconic memory is due to George Sperling (1960). The starting point for Sperling's work is the classical experiment on the determination of the memory-span. If a certain number of elements, letters, for example are presented using a tachistoscope it can be seen that the number of correct responses does not exceed 4 to 5 regardless of the number of letters presented. Provided that the number of elements presented remains below this limit, all the elements are reproduced correctly; if more than this are presented performance remains at this limit however many are presented. This limit is known as the memory-span.

The notion of whole report procedure excited a lot of interest at the end of the last century after Javal discovered the phenomenon of fixations and saccades : presentation by means of the tachistoscope became a way of simulating what happens during a fixation. Whole report procedure suggests the existence of a fixed limit to the sensory information which can be processed during a fixation. The idea that whole report procedure reveals a limitation to the quantity of sensory information processed in a presentation did not always seem to agree with subjects reporting the impression hat they had seen more elements than they could remember. This impression would suggest another interpretation of the procedure which was that it was a limitation of mnemonic order.

Sperling's main contribution has been to find a technique which has supplied a clear answer to the question we have just posed: this is the partial report procedure. In one of his experiments he presents a matrix of 3 lines of four letters each to his subject for 50msec. If the subject is asked to reproduce what he has seen, on average four answers are correct. Under partial report procedure a high-, medium- or low- pitched tone is produced at the same time as the presentation is over and the task given to the subject is to reproduce only the first, second or third line according to the pitch of the tone. A performance in the order of three correct responses was obtained under these conditions. Since the subject was unable to foretell what line he would be asked for to report, it must be acknowledged that immediately after the end of the presentation, the information necessary to recall the letters had to be available somewhere.

This led Sperling to the hypothesis that one form of presentation of the visual stimulation remains accessible for a short time after the presentation. It was Neisser in 1967 who proposed that this representation should be called an icon.

The reason why it is not possible to reproduce nine elements in the case of whole report procedure is that the iconic trace deteriorates while the first elements are reproduced and by the time it takes to have reproduced four or five elements there is no longer any iconic memory available.

The temporary decline of iconic memory has been studied by Sperling in another experiment where the index of selection was not present immediately after the disappearance of the matrix but at different intervals. The results show that in about one second the performance falls to the level which corresponds with whole report procedure.

Nature of coding in iconic memory

Two principal properties are generally attributed to the icon.

  1. It is a "pre-categorical" trace, i.e. a relatively literal copy of the pattern of the visual stimulation which has undergone very little perceptual treatment. In one of his experiments Sperling compared

two selection criteria for partial recall : selection on the basis of position (one of the three lines) as in the previous experiments, selection on the basis of a semantic category : give the letters or give the figures (in this experiment there was a mixture of letters and figures). With the latter criterion there was no observable superiority of partial over total recall. This result would suggest that information on placement is immediately available in the icon while information on belonging to one of the categories is not and probably will not become so until after perceptual treatment which, having to be carried out on all the elements, cancels the advantage of partial recall.

2. It is a "visual" trace. This statement is based on the finding that the factors which have an effect on the efficiency of iconic storage are generally events limited to the visual domain. Iconic stocking is influenced by factors such as intensity, duration, figure-ground contrast of the images presented, stimulations which precede and follow the stimulation-target. On the other hand, stimulation in other sensory modalities, or mental activities, would have little effect on the characteristics of the icon.

In 1960, Sperling performed an experiment using a matrix with three rows of three letters. Participants of the study were asked to look at the letters, for a brief period of time, and then recall them immediately afterwards. This technique, called free recall showed that participants were able to, on average, recall 4-5 letters of the 9 they were given. This however, was already generally accepted in the psychological community, but Sperling believed that all 9 letters were stored in the viewer's memory for a short period of time, but the memory failed so rapidly that only 4 or 5 could be recalled. Sperling called this iconic memory. Sperling showed this with his experiment of cued recall. This trial was similar to free recall, however, instead if allowing participants to recall any of the letters, it would allow them to view the same matrix for the same amount of time, and then hear a pitch corresponding to a different row in the matrix. The viewer was to recall the letters in that corresponding row. On average, viewers were able to recall more during cued recall trials than free recall.

Sperling built upon this experiment to then determine the amount of time before information was discarded from a person's memory. Using the same matrix, allowing viewers to see the matrix for the same amount of time, and still giving the pitches to cue the viewer which row to recall, Sperling added a twist, there would be a 5 milli-second delay after the letters disappeared before the cue would appear. The participants were unable to recall as many letters, thus showing that visual stimuli that are not added to short term memory are discarded less than 5 milliseconds of initial introduction. (It was later agreed upon that most visual icons are eliminated from memory before 250 milliseconds)

Iconic Memory

It can store a great deal of memory only for a very brief period of time. In human experimentation, the subject is made to sit and is asked to look at a screen, a computer screen normally. A dot is shown on the screen with blank white screen (do is located in the center) and the subject is asked to focus on the dot. Instead of dot, a set of letter is presented at that fixed point, where the subject is focusing. After a very brief period letters (stimuli) are removed. Subjects are asked to report items. Subjects normally reported only 4-5 items out of 12 letters.

The set of stimuli is following

Sperling's partial report procedure

An experiment was conducted by Sperling in the field of visual sensory memory. In that experiment the same array of letter was presented to the subjects but the subjects were asked to report the letters according to the cue (a beep in this case an auditory cue). After the array was turned off, a tune was sounded high, medium or low.

- High tone cue for reporting the top row
- Medium tone cue for middle row
- Low tone cue for Bottom row

Better performance

- High tone ¬ HBST (subject to report the first row)
- Medium tone ¬ AHMG (subject to report the second row)
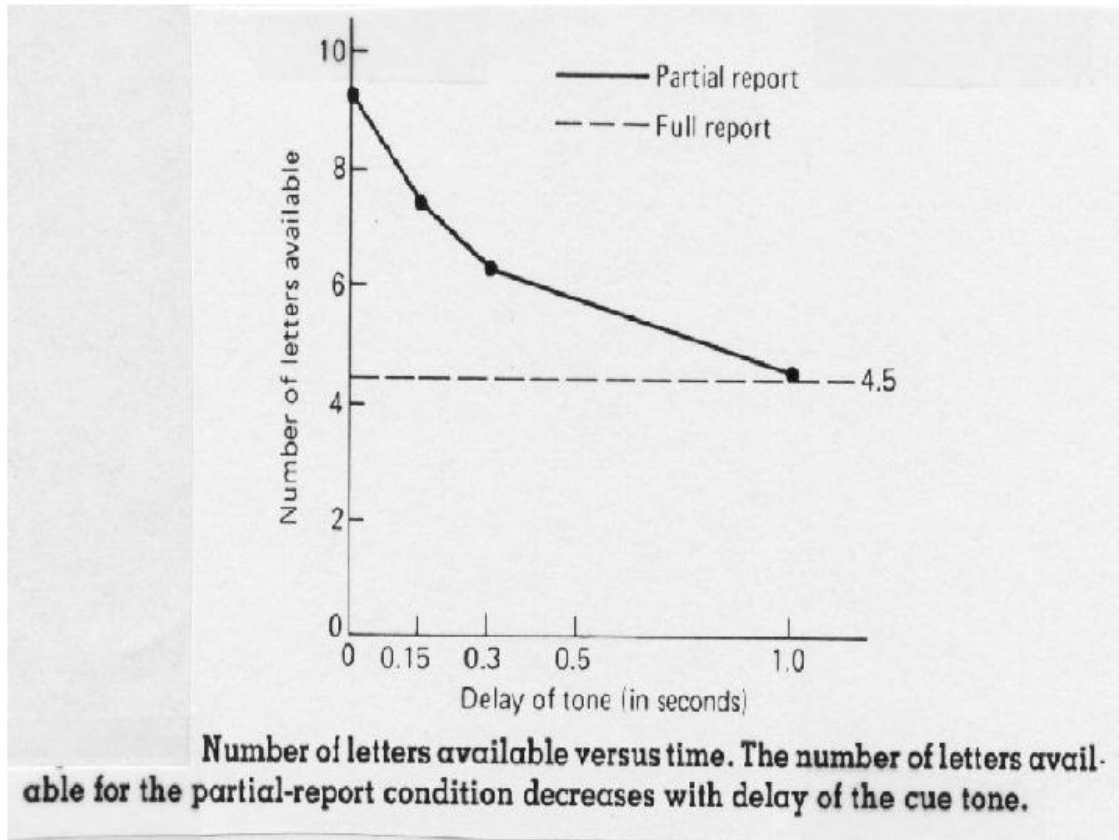- Low tone ¬ ELWJ (subject to report the third row)

Subjects reported at least 3 out of 4 letters, no matter whether it was on top row or middle row or bottom row. It was very significant. Because the subject did not know beforehand which row will be cued, they had to have three letters from each row available to them. So they had at least 9 out of 12 letters available in their visual memory Sperling's method was called the partial-report procedure.

The decay in visual memory

Sperling also varied the length of the delay between the offset of the display and the tone. As the delay increases to 1 second, subjects' performance decays back to the original whole-report level of 4 or 5 items. As the graph is showing. It indicates our visual memory lose most of the information in one second.

The decay in visual memory

Number of letters available versus time. The number of letters available for the partial-report condition decreases with delay of the cue tone.

The criticism on this experiment was that it is an artificial or laboratory phenomenon. It is not related to real life vision. Sperling's experiments indicate the existence of a brief visual sensory store- a memory that can effectively hold all the information in the visual display. We receive a lot of visual information from our surrounding but this all store in our visual sensory memory for a very brief period. That information receives attention, stored in our memory other is lost.

Sperling (1967) & Neisser (1967)

Sperling made another variation and after that array disappeared, he made visual field dark instead of white. This produced fascinating results and the retention power of the subjects was increased to 5 seconds. He found that when the post-exposure field was light the sensory information remained for only a second, but when the field was dark it remained for a full 5 seconds.

Light post-exposure field ¬ 1 second
Dark post-exposure field ¬ 5 seconds

Ulric Neisser wrote first cognitive psychology book in 1957. He devised a word icon. It is a brief visual memory revealed in these experiments. He devised the word iconic memory for short term visual memory.

According to Neisser, the visual memory is neither short term memory nor long term memory but it is very very short term memory and should b called as iconic memory. Without such a visual icon, perception would be much more difficult. Many stimuli are of very duration. In order to recognize them, the system needs some means of holding on to them for a short while until they can be analyzed.

Neisser also reported that if another display is given during that one second when you are already retaining an image it is like erasing or washing out the first and overwriting the second one. Almost all the information is held for a very brief period (1 second). It is quickly washed out after removal of stimulus unless attention is paid to it.

The sensory store is particularly visual in character and is sensitive to light. We cannot pinpoint where sensory memory is located in brain (this is a software level description not hardware level description). Sensory visual store is not physical but seems like physical phenomenon and is sensitive to light.

(c) With suitable examples, critically evaluate both primary and secondary reinforcements and bring out their role in establishing conditioning.

TOPIC: *Learning*
SUBTOPIC: *Concept and theories of learning*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

Primary reinforcers

A primary reinforcer, sometimes called an unconditioned reinforcer, is a stimulus that does not require pairing to function as a reinforcer and most likely has obtained this function through the evolution and its role in species' survival. Examples of primary reinforcers include sleep, food, air, water, and sex. Some primary reinforcers, such as certain drugs, may mimic the effects of other primary reinforcers. While these primary reinforcers are fairly stable through life and across individuals, the reinforcing value of different primary reinforcers varies due to multiple factors (e.g., genetics, experience). Thus, one person may prefer one type of food while another abhors it. Or one person may eat lots of food while another eats very little. So even though food is a primary reinforcer for both individuals, the value of food as a reinforcer differs between them.

Secondary reinforcers

A secondary reinforcer, sometimes called a conditioned reinforcer, is a stimulus or situation that has acquired its function as a reinforcer after pairing with a stimulus that functions as a reinforcer. This stimulus may be a primary reinforcer or another conditioned reinforcer (such as money). An example of a secondary reinforcer would be the sound from a clicker, as used in clicker training. The sound of the clicker has been associated with praise or treats, and subsequently, the sound of the clicker may function as a reinforcer. As with primary reinforcers, an organism can experience satiation and deprivation with secondary reinforcers.

Primary reinforcers are biological. Food, drink, and pleasure are the principal examples of primary reinforcers. But, most human reinforcers are secondary, or conditioned. Examples include money, grades in schools, and tokens.

Secondary reinforcers acquire their power via a history of association with primary reinforcers or other secondary reinforcers. For example, if I told you that dollars were no longer going to be used as money, then dollars would lose their power as a secondary reinforcer.

Here's an example of how a secondary reinforcer is established. Let's train a dog to sit. First we would introduce the discriminative stimulus, the word "sit." We could just say "sit" and when the dog sits, we would give it some food. The food would be the primary reinforcer. Immediately after we gave it the food we would say, "good dog." "Good dog" is our secondary reinforcer of praise. We would then repeat the above process many times. Gradually, we would give the food less often, but the dog would continue to sit when we told it to. The words "good dog" gradually became a secondary reinforcer.

Another example would be in a token economy. Many therapeutic settings use the concept of the token economy. Remember, a token is just an object that symbolizes some other thing. For example, poker chips are tokens for money. In New York City, subway tokens are pieces of metal that can be inserted into the turnstiles of the subway. Small debts can be paid off using tokens in New York because of the token's value

of one subway ride. However, attempting to pay off debts elsewhere using NYC subway tokens would not be acceptable.

In a token economy, people earn tokens for making certain responses; then those tokens can be cashed in for privileges, food, or drinks. For example, residents of an adolescent halfway house may earn tokens by making their beds, being on time to meals, not fighting, and so on. Then, being able to go to the movies on the weekend may require a certain number of tokens.

Primary reinforcers are almost always tangible. They usually consist of something the child can hold or feel but they always involve a direct desire. Examples include a favourite ball, tunnel, toy, video, or other things that arouse the senses like bubbles, tickles, hugs or squeezes, textures, or music. One of the most basic primary reinforcers is food. Food can be a reinforcer even when your child is not hungry, if it is a preferred snack. The strategy is to only give a very small amount of food after a set amount of successful responses or tasks. A favourite snack can go along way if managed appropriately. It is also important to not let that snack or object become over-indulged. In this case, too much of a good thing can create havoc. Students, especially those with autism, can become overstimulated by a highly preferred activity, object, or food. They may perseverate on that one thing and feel the need to have it all the time. They may also have negative behaviors when that reinforcer is then taken away. Or the opposite can occur, and the student may not find that item desirable anymore because he sees it all the time. In order to prevent this from happening, it is important to offer multiple reinforcers throughout the day and limit the use of any one highly preferred item. In another post, I will talk about how to develop a list of reinforcers that your child loves and how to rotate them or schedule them appropriately.

Secondary reinforcers, as mentioned above, are learned. They are intrinsic and rewarding on an internal level, giving the student a feeling or the anticipation of something that they eventually associate with an activity. For example, the reading of a bedtime story can be associated with feeling sleepy if it is always read at around the same time, in bed, just before sleep. Some other examples of secondary reinforcement include verbal praise, smiles, tokens, thumbs up, and clapping. For typical students, verbal praise is usually enough. Children realize that they are doing something good when they get excitement and smiles from the adults or peers around them. With children who lack social empathy and the ability to relate to others' feelings, this verbal praise needs to be paired with something else. If the child loves to be hugged or squeezed, you may want to pair verbal praise with a big hug to create a good, warm feeling. The student will pair this secondary reinforcer with this feeling if it is always presented together. Secondary reinforcers can also become the promise or anticipation of a bigger reward. This is set up with a token system. To better prepare students for secondary reinforcement, you can set up tokens that the student earns after successful tasks. After a set amount of tokens are earned, a bigger, more tangible reward can be given. For example, a student may need to earn 10 tokens before he gets to go on the computer to watch a favourite video or play a favourite game. In order to pair these tokens with something good, an immediate primary reinforcer may be used with the tokens, like a piece of food, each time a token is earned and then intermittently until tokens are accepted and become motivating enough. This teaches a child to also wait for gratification or something desirable. Waiting can be very difficult for children with special needs so token boards can set up a nice scenario to work on this skill.

3. Answer the following, each in not more than 400 words: **30x2=60**

(a) In which way IRT is an improvement over classical test theory? Compare the two approaches and critically evaluate Rasch's model of IRT.

TOPIC: *Research Methods*
SUBTOPIC: *Item Response Theory*
LEVEL: *Medium*
NATURE: *Fundamental*

REFERENCE:

Item Response Theory

In psychometrics, **item response theory (IRT)** also known as **latent trait theory**, **strong true score theory**, or **modern mental test theory**, is a paradigm for the design, analysis, and scoring of tests, questionnaires, and similar instruments measuring abilities, attitudes, or other variables. It is based on the application of related mathematical models to testing data. Because it is generally regarded as superior to classical test theory, it is the preferred method for the development of high-stakes tests such as the Graduate Record Examination (GRE) and Graduate Management Admission Test (GMAT).

The name *item response theory* is due to the focus of the theory on the item, as opposed to the test-level focus of classical test theory, by modeling the *response* of an examinee of given ability to each *item* in the test. The term *item* is used because many test questions are not actually questions; they might be multiple choice questions that have incorrect and correct responses, but are also commonly statements on questionnaires that allow respondents to indicate level of agreement (a rating or Likert scale), or patient symptoms scored as present/absent. IRT is based on the idea that the probability of a correct/keyed response to an item is a mathematical function of person and item parameters. There is but a single person parameter; it is called latent trait or ability; it may, for example, be a unidimensional representation of a person's intelligence or the strength of an attitude. Item parameters include difficulty (location), discrimination (slope or correlation), and pseudoguessing (lower asymptote).

The concept of the item response function was around before 1950. The pioneering work of IRT as a theory occurred during the 1950s and 1960s. Three of the pioneers were the Educational Testing Service psychometrician Frederic M. Lord, the Danish mathematician Georg Rasch, and Austrian sociologist Paul Lazarsfeld, who pursued parallel research independently. Key figures who furthered the progress of IRT include Benjamin Drake Wright and David Andrich. IRT did not become widely used until the late 1970s and 1980s, when personal computers gave many researchers access to the computing power necessary for IRT.

Among other things, the purpose of IRT is to provide a framework for evaluating how well assessments work, and how well individual items on assessments work. The most common application of IRT is in education, where psychometricians use it for developing and refining exams, maintaining banks of items for exams, and equating for the difficulties of successive versions of exams (for example, to allow comparisons between results over time).

IRT models are often referred to as *latent trait models*. The term *latent* is used to emphasize that discrete item responses are taken to be *observable manifestations* of hypothesized traits, constructs, or attributes, not directly observed, but which must be inferred from the manifest responses. Latent trait models were developed in the field of sociology, but are virtually identical to IRT models.

IRT is generally regarded as an improvement over classical test theory (CTT). For tasks that can be accomplished using CTT, IRT generally brings greater flexibility and provides more sophisticated information. Some applications, such as computerized adaptive testing, are enabled by IRT and cannot reasonably be performed using only classical test theory. Another advantage of IRT over CTT is that the more sophisticated information IRT provides allows a researcher to improve the reliability of an assessment.

IRT entails three assumptions:

1. A unidimensional trait denoted by θ;
2. Local independence of items;
3. The response of a person to an item can be modeled by a mathematical *item response function* (IRF).

The trait is further assumed to be measurable on a scale (the mere existence of a test assumes this), typically set to a standard scale with a mean of 0.0 and a standard deviation of 1.0. 'Local independence' means that items are not related except for the fact that they measure the same trait, which is equivalent to the assumption of unidimensionality, but presented separately because multidimensionality can be caused by other issues. The topic of dimensionality is often investigated with factor analysis, while the IRF is the basic building block of IRT and is the center of much of the research and literature.

Classical test theory (CTT and item response theory (IRT) are widely perceived as representing two very different measurement frameworks. However, few studies have empirically examined the similarities and differences in the parameters estimated using the two frameworks. Prior to exploring this issue in some detail, some readers may appreciate a brief review of related theories. Additional detail is provided elsewhere (cf. Crocker & Algina, 1986; McKinley & Mills, 1989).

Brief Review of CTT and IRT

Although CTT has served the measurement community for most of this century, IRT has witnessed an exponential growth in recent decades. The major advantage of CTT are its relatively weak theoretical assumptions, which make CTT easy to apply in many testing situations (Hambleton & Jones, 1993). Relatively weak theoretical assumptions not only characterize CTT but also its extensions (e.g., generalizability theory). Although CTT's major focus is on test-level information, item statistics (i.e., item difficulty and item discrimination) are also an important part of the CTT model.

At the item level, the CTT model is relatively simple. CTT does not invoke a complex theoretical model to relate an examinee's ability to success on a particular item. Instead, CTT collectively considers a pool of examinees and empirically examines their success rate on an item (assuming it is dichotomously scored). This success rate of a particular pool of examinees on an item, well known as the p value of the item, is used as the index for the item difficulty (actually, it is an inverse indicator of item difficulty, with higher value indicating an easier item). The ability of an item to discriminate between higher ability examinees and lower ability examinees is known as item discrimination, which is often expressed statistically as the Pearson product-moment correlation coefficient between the scores on the item (e.g., 0 and 1 on an item scored right-wrong) and the scores on the total test. When an item is dichotomously scored, this estimate is often computed as a point-biserial correlation coefficient.

The major limitation of CTT can be summarized as circular dependency: (a) The person statistic (i.e., observed score) is (item) sample dependent, and (b) the item statistics (i.e., item difficulty and item discrimination) are (examinee) sample dependent. This circular dependency poses some theoretical difficulties in CTT's application in some measurement situations (e.g., test equating, computerized adaptive testing).

Despite the theoretical weakness of CTT in terms of its circular dependency of item and person statistics, measurement experts have worked out practical solutions within the framework of CTT for some otherwise difficult measurement problems. For example, test equating can be accomplished empirically within the CTT framework (e.g., equipercentile equating). Similarly, empirical approaches have been proposed to accomplish item-invariant measurement (e.g., Thurstone absolute scaling) (Englehard, 1990). It is fair to say that, to a great extent, although there are some issues that may not have been addressed theoretically within the CTT framework, many have been addressed through ad hoc empirical procedures.

IRT, on the other hand, is more theory grounded and models the probabilistic distribution of examinees' success at the item level. As its name indicates, IRT primarily focuses on the item-level information in contrast to the CTT's primary focus on test-level information. The IRT framework encompasses a group of models, and the applicability of each model in a particular situation depends on the nature of the test items and the viability of different theoretical assumptions about the test items. For test items that are dichotomously scored, there are three IRT models, known as three-, two-, and one-parameter IRT models.

Although the one-parameter model is the simplest of the three models, it may be better to start from the most complex, the three-parameter IRT model; the reason for this sequence of discussion will soon become obvious. The IRT three-parameter model takes the following form:

(1) [MATHEMATICAL EXPRESSION NOT REPRODUCIBLE IN ASCII]

where c_i is the guessing factor, a_i is the item discrimination parameter commonly known as item slope, b_i is the item difficulty parameter commonly known as the item location parameter, D is

19

an arbitrary constant (normally, D = 1.7), and [Theta] is the ability level of a particular examinee. The item location parameter is on the same scale of ability, [Theta], and takes the value of [Theta] at the point at which an examinee with the ability-level [Theta] has a 50/50 probability of answering the item correctly. The item discrimination parameter is the slope of the tangent line of the item characteristic curve at the point of the location parameter.

When the guessing factor is assumed or constrained to be zero ([c.sub.i] = 0), the three-parameter model is reduced to the two-parameter model for which only item location and item slope parameters need to be estimated:

(2) [MATHEMATICAL EXPRESSION NOT REPRODUCIBLE IN ASCII]

If another restriction is imposed that stipulates that all items have equal and fixed discrimination, then [a.sub.i] becomes a, a constant rather than a variable, and as such, this parameter does not require estimation, and the IRT model is further reduced to

(3) [MATHEMATICAL EXPRESSION NOT REPRODUCIBLE IN ASCII]

So, for the one-parameter IRT model, constraints have been imposed on two of the three possible item parameters, and item difficulty remains the only item parameter that needs to be estimated. This one-parameter model is often known as the **Rasch** model, named after the researcher who did pioneer work in the area. It is clear from the discussion that the three-parameter model is the most general model, and the other two IRT models (two- and one-parameter models) can be considered as models nested or subsumed under the three-parameter model.

Theoretically, IRT overcomes the major weakness of CTT, that is, the circular dependency of CTT's item/person statistics. As a result, in theory, IRT models produce item statistics independent of examinee samples and person statistics independent of the particular set of items administered. This invariance property of item and person statistics of IRT has been illustrated theoretically (Hambleton & Swaminathan, 1985; Hambleton, Swaminathan, & Rogers, 1991) and has been widely accepted within the measurement community.

The invariance property of IRT model parameters makes it theoretically possible to solve some important measurement problems that have been difficult to handle within the CTT framework, such as those encountered in test equating and computerized adaptive testing (Hambleton et al., 1991). However, as the cornerstone of IRT, the importance of the invariance property of IRT model parameters cannot be overstated, because, without this crucial property, the complexity of IRT models can hardly be justified on either theoretical or practical grounds.

Because IRT differs considerably from CTT in theory, and commands some crucial theoretical advantages over CTT, it is reasonable to expect that there would be appreciable differences between the IRT- and CTT-based item and person statistics. Theoretically, such relationships are not entirely clear, except that the two types of statistics should be monotonically related under certain conditions (Crocker & Algina, 1986; Lord, 1980). But such relationships have rarely been empirically investigated, and, as a result, they are largely unknown.

The empirical studies available in this area have primarily focused on the application of the two methods in test equating (e.g., Becker & Forsyth, 1992; Harris, 1991). With regard to test equating, Hambleton et al. (199 1) suggested that, theoretically, the invariance property of the IRT item statistics obviated the need of equating tests; instead, it is (linear) scaling, rather than equating, that is necessary within the framework of IRT. The discussion implies that IRT models handle equating tasks better than the CTT equating approaches. The empirical studies in this area, however, provide a mixed picture, with some indicating the superiority of IRT approaches (e.g., Peterson, Cook, & Stocking, 1983), some suggesting better results from CTT ad hoc approaches (e.g., Clemans, 1993; Kolen, 1981; Skaggs & Lissitz, 1986a), and still some finding that both CTT and IRT equating methods produce very comparable results (Skaggs & Lissitz, 1988). The

mixed picture has prompted some researchers to suggest that it might be unrealistic to expect one method to provide the best equating results for all types of tests (e.g., Skaggs & Lissitz, 1986b).

A literature search revealed only one study that empirically examined the comparability of IRT-based and CTT-based item and person statistics. Lawson (1991) compared IRT-based (one-parameter **Rasch** model) and CTT-based item and person statistics for three different data sets, and showed exceptionally strong relationships between the IRT- and CTF-based item and person statistics. The results of the study, although the study was based on somewhat small data sets and only examined the most restrictive one-parameter IRT model, suggest that information from the two approaches about items and examinees might be very much the same.

Similarly, the invariance property of IRT item/person parameters has been little explored empirically, although invariance has been illustrated theoretically (e.g., Hambleton & Swaminathan, 1985; Rudner, 1983). However, Miller and Linn (1988), using an extant large data set, did report the results of a study examining the variations of item characteristic functions in the context of instructional coverage variations. They reported relatively large differences in item curve responses, suggesting lack of invariance of IRT item parameters. Lack of invariance was also reported by Cook, Eignor, and Taft (1988) for both CTT- and IRT-based item difficulty estimates.

Given the limited number of empirical studies directly or indirectly addressing the invariance issue, there is an obvious lack of systematic investigation about the absolute invariance of the item and person statistics obtained from either CTF or IRT frameworks and a lack of studies that empirically compare the relative invariance of item and person statistics obtained from CTT versus those from IRT. The major criticism for CTT is its inability to produce item/person statistics that would be invariant across examinee/item samples. This criticism has been the major impetus for the development of IRT models and for the exponential growth of IRT research and applications in the recent decades.

It is somewhat surprising that empirical studies examining and/or comparing the invariance characteristics of item statistics from the two measurement frameworks are so scarce. It appears that the superiority of IRT over CTT in this regard has been taken for granted by the measurement community, and no empirical scrutiny has been deemed necessary. The empirical silence on this issue seems to be an anomaly. This lack of empirical investigation has prompted some researchers to state that item response modeling has been too focused on mathematical elaboration at the expense of empirical exploration (Goldstein & Wood, 1989).

A comparison of classical and item response theories

Classical test theory (CTT) and IRT are largely concerned with the same problems but are different bodies of theory and entail different methods. Although the two paradigms are generally consistent and complementary, there are a number of points of difference:

- IRT makes stronger assumptions than CTT and in many cases provides correspondingly stronger findings; primarily, characterizations of error. Of course, these results only hold when the assumptions of the IRT models are actually met.
- Although CTT results have allowed important practical results, the model-based nature of IRT affords many advantages over analogous CTT findings.
- CTT test scoring procedures have the advantage of being simple to compute (and to explain) whereas IRT scoring generally requires relatively complex estimation procedures.
- IRT provides several improvements in scaling items and people. The specifics depend upon the IRT model, but most models scale the difficulty of items and the ability of people on the same metric. Thus the difficulty of an item and the ability of a person can be meaningfully compared.
- Another improvement provided by IRT is that the parameters of IRT models are generally not sample- or test-dependent whereas true-score is defined in CTT in the context of a specific test. Thus IRT provides significantly greater flexibility in situations where different samples or test forms are used. These IRT findings are foundational for computerized adaptive testing.

It is worth also mentioning some specific similarities between CTT and IRT which help to understand the correspondence between concepts. First, Lord showed that under the assumption that θ is normally distributed, discrimination in the 2PL model is approximately a monotonic function of the point-biserial correlation. In particular:

$$a_i \cong \frac{\rho_{it}}{\sqrt{1 - \rho_{it}^2}}$$

where ρ_{it} is the point biserial correlation of item *i*. Thus, if the assumption holds, where there is a higher discrimination there will generally be a higher point-biserial correlation.

Another similarity is that while IRT provides for a standard error of each estimate and an information function, it is also possible to obtain an index for a test as a whole which is directly analogous to Cronbach's alpha, called the *separation index*. To do so, it is necessary to begin with a decomposition of an IRT estimate into a true location and error, analogous to decomposition of an observed score into a true score and error in CTT. Let

$$\hat{\theta} = \theta + \varepsilon$$

where θ is the true location, and ϵ is the error association with an estimate. Then $\text{SE}(\theta)$ is an estimate of the standard deviation of ϵ for person with a given weighted score and the separation index is obtained as follows

$$R_\theta = \frac{\text{var}[\theta]}{\text{var}[\hat{\theta}]} = \frac{\text{var}[\hat{\theta}] - \text{var}[\epsilon]}{\text{var}[\hat{\theta}]}$$

where the mean squared standard error of person estimate gives an estimate of the variance of the errors, ϵ_n, across persons. The standard errors are normally produced as a by-product of the estimation process. The separation index is typically very close in value to Cronbach's alpha.

IRT is sometimes called *strong true score theory* or *modern mental test theory* because it is a more recent body of theory and makes more explicit the hypotheses that are implicit within CTT.

The Rasch model

The Rasch model is often considered to be the 1PL IRT model. However, proponents of Rasch modeling prefer to view it as a completely different approach to conceptualizing the relationship between data and the theory. Like other statistical modeling approaches, IRT emphasizes the primacy of the fit of a model to observed data, while the Rasch model emphasizes the primacy of the requirements for fundamental measurement, with adequate data-model fit being an important but secondary requirement to be met before a test or research instrument can be claimed to measure a trait. Operationally, this means that the IRT approaches include additional model parameters to reflect the patterns observed in the data (e.g., allowing items to vary in their correlation with the latent trait), whereas the Rasch approach requires both the data fit the Rasch model and that test items and examinees confirm to the model, before claims regarding the presence of a latent trait can be considered valid. Therefore, under Rasch models, misfitting responses require diagnosis of the reason for the misfit, and may be excluded from the data set if substantive explanations can be made that they do not address the latent trait. Thus, the Rasch approach can be seen to be a confirmatory approach, as opposed to exploratory approaches that attempt to model the observed data. As in any confirmatory analysis, care must be taken to avoid confirmation bias.

The presence or absence of a guessing or pseudo-chance parameter is a major and sometimes controversial distinction. The IRT approach includes a left asymptote parameter to account for guessing in multiple

choice examinations, while the Rasch model does not because it is assumed that guessing adds randomly distributed noise to the data. As the noise is randomly distributed, it is assumed that, provided sufficient items are tested, the rank-ordering of persons along the latent trait by raw score will not change, but will simply undergo a linear rescaling. Three-parameter IRT, by contrast, achieves data-model fit by selecting a model that fits the data, at the expense of sacrificing specific objectivity.

In practice, the Rasch model has at least two principal advantages in comparison to the IRT approach. The first advantage is the primacy of Rasch's specific requirements, which (when met) provides *fundamental* person-free measurement (where persons and items can be mapped onto the same invariant scale). Another advantage of the Rasch approach is that estimation of parameters is more straightforward in Rasch models due to the presence of sufficient statistics, which in this application means a one-to-one mapping of raw number-correct scores to Rasch θ estimates.

For Details:

[PDF File Link: Item Response Theory and Classical Test Theory - An Empirical Comparison](#)

(b) Describe Broadbent's filter theory of attention. Why this elegant theory turned out to be inadequate? Comment upon the alternate approaches.

TOPIC: *Sensation, Attention and Perception*
SUBTOPIC: *Factors Influencing Attention*
LEVEL: *Medium*
NATURE: *Fundamental*

REFERENCE:
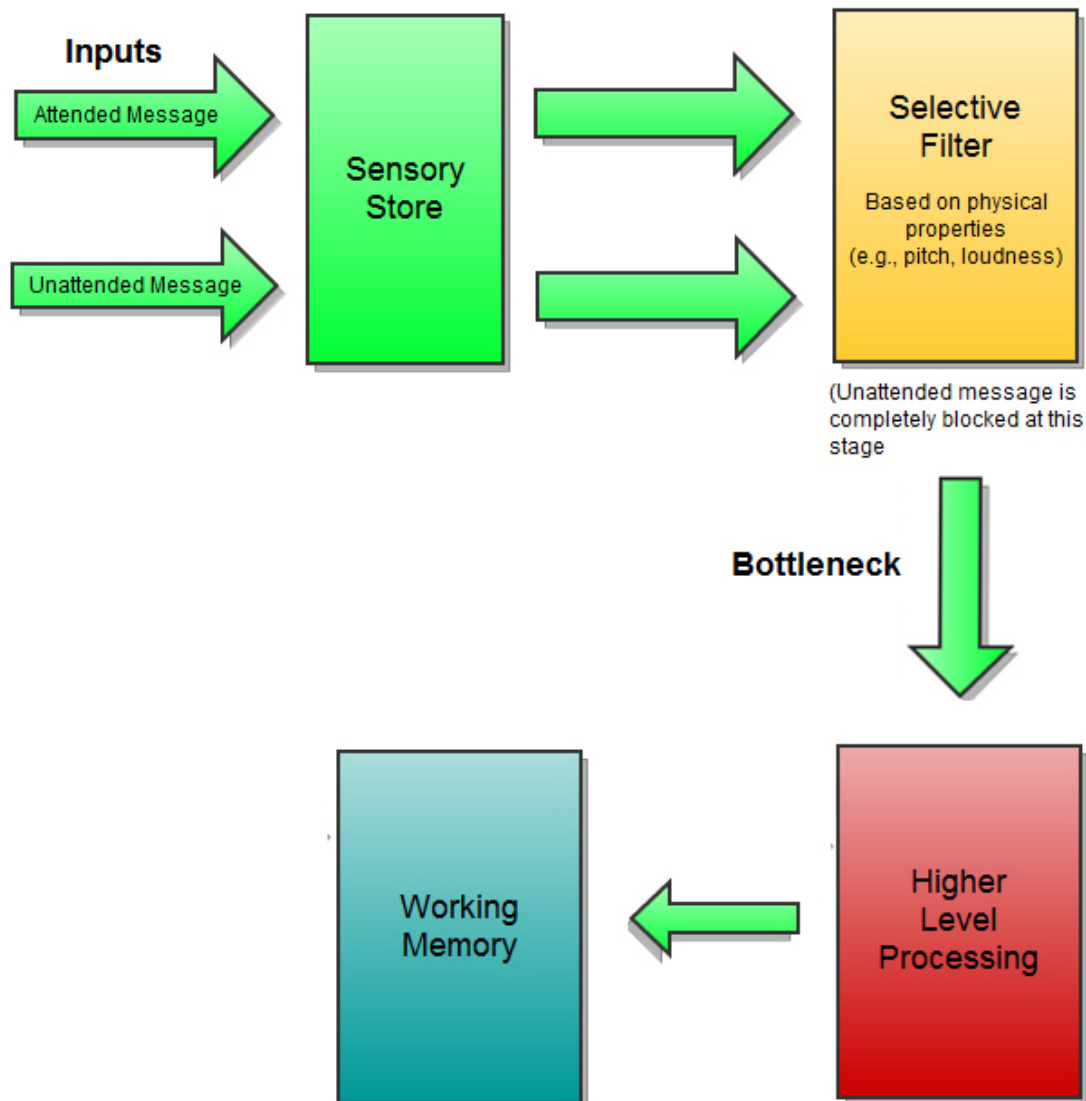
Broadbent's filter model of attention

Broadbent's filter model is an early selection theory of attention.

Description

Donald Broadbent developed the filter model as an extension of William James' multi-storage paradigm. Broadbent proposed the notion that a filter acts as a buffer on incoming sensory information to select what information gains conscious awareness. The attended information will pass through the filter, while unattended information will be completely blocked and ignored. The filter acts on stimuli solely on their physical characteristics, such as location, loudness, and pitch.

During World War II the rapid development of machinery did not arise without complications. It was common for radar operators to have difficulties communicating with several pilots at once, as all of their voices were broadcast over one loud speaker. This overloading of perceptual input fuelled Broadbent's curiosity of how stimuli capture our attentional resources.

23

Broadbent's Filter Model

In the filter model, initial processing of stimuli occurs pre-attentively on the basis of their physical features, and is housed in a temporary sensory store. Unlike the physical properties, Broadbent believed semantic features, due to their complexity, would impose a limited capacity on the temporary storehouse of incoming stimuli. For this reason, he postulated a filter then acts on the stimuli, to determine what will be processed further and filter out irrelevant stimuli. Information selected to pass through the filter is then available for short-term memory and manipulation of the selected information, prior to storage in long-term memory.

The development of the filter model was the first theoretical account relating psychological phenomena to information processing concepts of mathematics and computer science. As so, Broadbent provided a computer metaphor in which information-processing at the micro level acted in series, while at the macro level it operated in a parallel fashion. As attention can be directed by physical properties or by an organism's drives, this reveals a parallel processing manner at the macro level, while still processing information semantically at a micro level. Further, goal-directed behavior requires attention to be controlled; hence a high degree of selectivity is put forth in the information-processing stream. When

24

developing his model, Broadbent emphasized the splitting of incoming stimuli to attended or unattended channels. Channel selection is guided through attention. If one is attempting to attend to a stimulus based on their current goals, they will employ voluntary attention; whereas if a sensory event catches one's attention, reflexive attention will be employed. During his experimentation, Broadbent made use of the dichotic listening test. This task has been used extensively to test numerous psychological phenomena such as response times of specific auditory information, as well as testing for attended and unattended information presented to a participant. It is widely used as it is a non-invasive method of testing cerebral dominance. In a typical dichotic listening task, the participant is wearing a headphone, which will have differing stimuli presented in each ear piece. The participant is instructed to attend (attended channel) the information coming from one of the ear pieces and neglect (unattended channel) the information presented from the other. Following the listening period, the participants are tested on whether they recall any information presented in the unattended channel. Early research using dichotic listening tasks provided empirical evidence of participants' ability to correctly recall information to the attended channel, and poor recalling in the unattended channel. Caution must be taken when considering the results of dichotomous listening tests as the majority of people have a right ear advantage for verbal stimuli. This means that information presented to the right ear has a better chance of gaining conscious awareness than information presented to the left ear.

[Why this elegant theory turned out to be inadequate?]

Modern views of Broadbent's model

As psychological research has improved immensely since Broadbent's time, more sophisticated measures indicate that we do have an attentional filter, though it is integrated into a broader cognitive system. This system compensates for the controversies of limited parallel processing in Broadbent's original findings. A major component of the system entails sensory memory, which is broken down into iconic memory and echoic memory. The aforementioned represent visual and auditory memory respectively, which function preattentively. Given the existence of such a preattentive memory store makes it possible for preattentive stimuli to work in a serial manner. Research on iconic memory has provided a visual hierarchy of the visual system, which indicates specific neurons are activated before stimulus recognition, supporting Broadbent's theory of preattentive processing. Additionally, research has shown that physical features of a stimulus guide attentional selection. It has been found consistently that observers correctly separate relevant from irrelevant stimuli due to physical rather than semantic features, indicating selection channels are heavily influenced by physical features. According to the modality effect, echoic memory has an advantage over iconic memory. Research has shown that the auditory system is more apt to objective interpretation than the visual system, as auditory organization is less ambiguous than the visual organization. This indicates that auditory information is first processed for its physical features, and then combined with visual information features. Moreover, allocation of attention is a product of both voluntary and reflexive attention. Goals and behaviors drive attention but may be influenced by an exogenous stimulus of particular stimulus strength, which varies by situation. Such research evidence confirms Broadbent's notion of voluntary attentional mechanisms. More recent research finds that Broadbent's model neglected to address the time requirements of shifting attention. However, he did distinguish that internal and external stimuli can cause shifts of attention, though he did not consider that internally and externally driven shifts of attention may have differing time courses.

[Comment upon the alternate approaches.]

Attention

Attention is commonly understood as the ability to attend to some things while ignoring others. Attention is controllable, selective, and limited. It is the progression by which external stimuli form internal representations that gain conscious awareness. Attention is part of nearly every waking moment for humans, as it is the focusing of one's thoughts. Selective attention utilizes cognitive processes to focus on relevant targets on input, thoughts or actions while neglecting irrelevant sources of input. This is the basis for how we attend to specific stimuli. Voluntary attention, otherwise known as top-down attention, is the aspect over which we have control, enabling us to act in a goal-directed manner. In contrast, reflexive

attention is driven by exogenous stimuli redirecting our current focus of attention to a new stimulus, thus it is a bottom-up influence. These two divisions of attention are continuously competing to be the momentary foci of attention. Selection models of attention theorize how specific stimuli gain our awareness. Early selection models emphasize physical features of stimuli are attended to, while late selection models argue that semantic features are what determine our current focus of attention.

Early selection models of attention

Psychological research regarding attention is largely governed by the processes of which stimuli are selected to enter our conscious awareness. Selection models are utilized by researchers to propose when stimulus information is attended to. Early selection models state that a stimulus is selected to be attended to prior to coming in contact with the filter. A filter can be regarded as the selector of relevant information. Information which is relevant to a particular situation will be attended to, while irrelevant attention is then unattended to. Broadbent supported this early selection theory. The basic idea proposes that perception of the stimulus is not required prior to selecting its relevance.

Late selection models of attention

Late selection models are those which state that incoming information is not selected for by a filter specifically, but later in the information processing sequence. With these models, both attended and unattended information passes through the filter. The filter merely acts as an information attenuator, as in it intensifies the pertinent information, and attenuates the intensity of the stimuli deemed to be unimportant . Information inputs are processed equivalently, until semantic encoding and analysis can be performed. This notion implies that internal decisions of stimuli relevance must be made, before allowing it to gain conscious awareness.

Attenuation model of attention

Anne Treisman, a graduate student of Broadbent's, was not fully convinced by the notion of a filter performing decisions as to what stimuli gain conscious awareness. She proposed an alternative mechanism, the attenuation theory, in which the filter acts as an attenuator of information, either increasing or decreasing attentional capacities towards it. This slight modification has the unattended channel passing through all processing stages, only weakened rather than completely blocked. As the unattended channel includes weakly attended to information, to gain conscious awareness this information must surpass a threshold, which Treisman believed was determined by the words' meaning. Important words (such as one's name) would have a low threshold, to easily gain awareness, whereas unimportant words (such as "lamp") would have a higher threshold to prevent them from gaining awareness inappropriately. In this way, the threshold for each word acts as a filtering mechanism, relying on semantic features.

Memory selection model of attention

Deutsch, Deutsch and Norman were not fully convinced by Broadbent's selection criteria based solely on physical features of a stimulus. The cocktail party effect influenced researchers to look further than physical selection features, to semantic selecting features. The cocktail party effect is an example of how unattended information can gain one's awareness. Suppose you were at a social gathering having a conversation with some friends, when you hear someone mention your name and it grasps your attention. This unattended-to information somehow gained your awareness. This fuelled the development of the memory selection model, which shares the same basic principle of early selection models that stimulus features are selected via their physical properties. Attended and unattended information passes through the filter, to a second stage of selection on the basis of semantic characteristics or message content. Items which are selected are incorporated into short-term memory and awareness. The second selection mechanism, rather than the filter, decides what information gains our awareness.

Multimode model of attention

Additional research proposes the notion of a moveable filter. The multimode theory of attention combines physical and semantic inputs into one theory. Within this model, attention is assumed to be flexible, allowing different depths of perceptual analysis. Which feature gathers awareness is dependent upon the person's needs at the time. Switching from physical and semantic features as a basis for selection yields costs and benefits. Stimulus information will be attended to via an early selection through sensory analysis, then as it increases in complexity, semantic analysis is involved, compensating for attention's limited capacity. Shifting from early to late selection models reduces the significance of stimuli rendering one's attention,[clarification needed] though it increases breadth of attention. Research has found that semantic selection requires a greater attentional resources than physical selection.

Capacity model of attention

Daniel Kahneman took a different approach to describing attention, by describing its division, rather than selection mechanisms. He describes attention as a resource in which energy or mental effort is required. Mental effort is used while engaging in performing any mental task, and the greater the complexity, the greater the effort needed to solve a task. Kahneman believes there are three basic conditions which needed to be met for proper completion of a task. By combining total attentional capacity, momentary mental effort, and appropriate allocation policy of the attentional capacity, a person will exert enough mental effort to overcome mental tasks. The key component is allocating enough attention, as a resource, to the task at hand. Kahneman also noted that arousal influences the total attentional capacity in any given situation. In addition, his model incorporates the ideas of voluntary and reflexive attention, which affect allocation policy. In order to direct attention appropriately, one must attend to relevant information, while neglecting irrelevant information to prevent becoming distracted. This mental effort theory proposed by Kahneman provides an overview of the influences and interdependencies of attention allocation, which is meant to supplement attention selection models.

4. Answer the following, each in not more than 250 words: 20x3=60

(a) What is McCrary-Hunter 'invariance hypothesis'? Discuss the shape and characteristics of serial position error curve in terms of 'invariance hypothesis'.

TOPIC: *Memory*
SUBTOPIC: *Encoding and remembering*
LEVEL: *Difficult*
NATURE: *Fundamental*

REFERENCE:

Hunter-McCrary Law

An evidence for scale-independence in memory is the fact that the proportion of errors produced for each serial position during serial learning remains constant even in the face of considerable variations in degree of learning, IPI, time between trials, familiarity or meaningfulness of the material to be remembered, or individual differences in the learners (Braun & Heymann, 1958; McCrary & Hunter, 1953); this is the **Hunter-McCrary Law**. Here the shape of the error distribution in serial learning provides the observer with no evidence about the absolute level of performance.

The Shape and Characteristics of Serial Position Error Curve in Terms of 'Invariance Hypothesis'

It is characteristic of scientific laws to hold over a wide range of temporal, spatial, or physical scales. We would be surprised, for example, if Newton's gravitational law, that the attractive force between two objects is inversely proportional to the squared distance between them, held for 1 mg objects but not 1 g or 1 kg objects, or for distances of centimeters but not distances of meters. Scientific principles are generally expected to hold at a wide range of scales (e.g. Barenblatt, 1996). Although laws may break down at extremes (e.g. for very small subatomic weights), the default search is for descriptions that are universal in that they apply over as wide a range of temporal and spatial scales as possible.

In developing models of human memory, however, it is widely assumed that different principles apply over different (short and long) time scales. Here we attempt to outline retrieval principles of human memory over both short and long time scales. We also relate time-invariant retrieval mechanisms to the mechanisms involved in identification of simple perceptual stimuli spread over either a narrow or broad psychophysical range. We term the resulting model SIMPLE (for Scale-Invariant Memory, Perception, and Learning). The model assumes that the confusability of

items in memory is given partly by the ratios of the items' temporal distances from the point of recall, and this emphasis on temporal ratios rather than absolute temporal durations gives the model its scale-invariant properties.

We begin the paper with a brief review of serial position effects and scale-invariant effects in memory and absolute identification. The subsequent theoretical development of the model can be summarized as follows. Early attempts to provide a unified account of serial position effects in absolute identification and memory tasks (Murdock, 1960) had difficulty accounting for isolation effects in absolute identification (Bower, 1971). We show how a simplified exemplar theory (Nosofsky, 1986) can shed light on the problematic isolation effects in absolute identification while also capturing the scale-invariant serial position effects observed in the same tasks. The model, SIMPLE, is applied to a novel experimental analog of primacy and recency effects in absolute identification. This paves the way for a return to Murdock's program of linking accounts of serial position effects in memory and identification paradigms. We then show that SIMPLE captures some scale-invariant and serial position effects in serial recall and free recall if it is assumed that (a) episodic memories are arrayed along a dimension representing temporal distance from the point of retrieval, (b) the retrievability of an item is inversely proportional to its summed confusability with other items in memory, and (c) the confusability of items along a temporal dimension is given by the ratio of the temporal distances of those items at the time of recall.

The model is used to account for many (although not all) data that have previously been assumed to reflect trace decay or the operation of a separate short-term memory, even though the model assumes that the same interference-based forgetting principles apply at all time scales. Thus the approach marks a return to the search for unitary principles of memory (Keppel & Underwood, 1962; Melton, 1963). The resulting model has affinities with a number of previous theoretical approaches, brings them together, resolves some of the difficulties, and argues that the combined approach can help to provide a unified perspective on a range of memory phenomena. The model can be viewed variously as: (a) an extension of Murdock's (1960) distinctiveness theory that accommodates time-based and local neighborhood effects, (b) a generalization and extension of early temporal discriminability and ratio models of memory (e.g. Baddeley, 1976; Bjork & Whitten, 1974; Crowder, 1976; Glenberg & Swanson, 1986), (c) the addition of a temporal dimension to a simplified exemplar model (Nosofsky, 1986, 1992), (d) the extension into the domain of temporal memory of the Feature Model's use of the Luce choice model for cue-driven recall (Nairne, 1990), (e) an extension of Neath's (1993a, 1993b) temporal distinctiveness model to allow isolation effects and primacy effects to be accounted for, or (f) a more analytic and abstract version of recent oscillator-based and contextual overlap models of memory for temporal order (e.g. Brown, Preece, & Hulme, 2000). More generally, following Gallistel (1990), the model places time and temporal interference at the heart of memory and relates memory retrieval to perceptual discriminability.

BASIC ISSUES

Serial Position Curves

Serial position effects have been seen as evidence for distinctions between long-term and short-term memory, as evidence for the greater distinctiveness of end-series items, and as evidence for a relation between serial recall and absolute identification. They are therefore central to the concerns of the present paper. Bowed serial position curves, showing reduced memory for mid-series items, are obtained in free recall, serial learning, probed serial recall, location memory, and some recognition memory tasks (for reviews see Crowder, 1976; Lansdale, 1998, McGeoch & Irion, 1952; Murdock, 1974). Serial position effects are also observed in absolute identification experiments in which items must be identified on the basis of their position along a single dimension such as frequency or amplitude (Attneave, 1950; Berliner, Durlach,

& Braida, 1977; Braida &Durlach, 1972; Durlach & Braida 1969; Kornbrot, 1978; Lacouture, 1997; Murdock, 1960; Woodworth & Thorndike, 1900). Serial position effects are also found in retrieval from long-term memory (Baddeley & Hitch, 1977; Bjork & Whitten, 1974; Glenberg, Bradley, Kraus, & Renzaglia, 1983; Healy, Havas, & Parker, 2000; Healy & Parker, 2001; Nairne, 1991; Pinto & Baddeley, 1991; Roediger & Crowder, 1976; Watkins, Neath, & Sechler, 1989) and in absolute identification experiments whether the stimuli to be identified span a narrow or wide psychophysical range (Lacouture, 1997; this paper). Many separate and incompatible theoretical accounts of these serial position effects are evident in the literature. Can a unified explanation be provided for these ubiquitous effects? Murdock (1960) answered affirmatively for the cases of identification and simple serial recall, suggesting that in both identification and recall items at the end of a series benefit through their greater relative distinctiveness. Accounts of serial position effects have subsequently diverged and proliferated. For example, in absolute identification, serial position effects have been attributed to response mapping (Lacouture & Marley, 1991, 1995), end anchor effects (Berliner & Durlach, 1973), response bias in combination with choice model application (Lansdale, 1998; Nosofsky, 1985), selective attention to end items (Luce, Green, & Weber, 1976), neighborhood density (Krumhansl, 1978), increased memory variance for mid-list items (Nosofsky, 1997), error-minimizing decision strategies (Balakrishnan, 1997), or dynamic criterion setting and response availability within a Thurstonian framework (Treisman, 1985; Treisman & Williams, 1984).
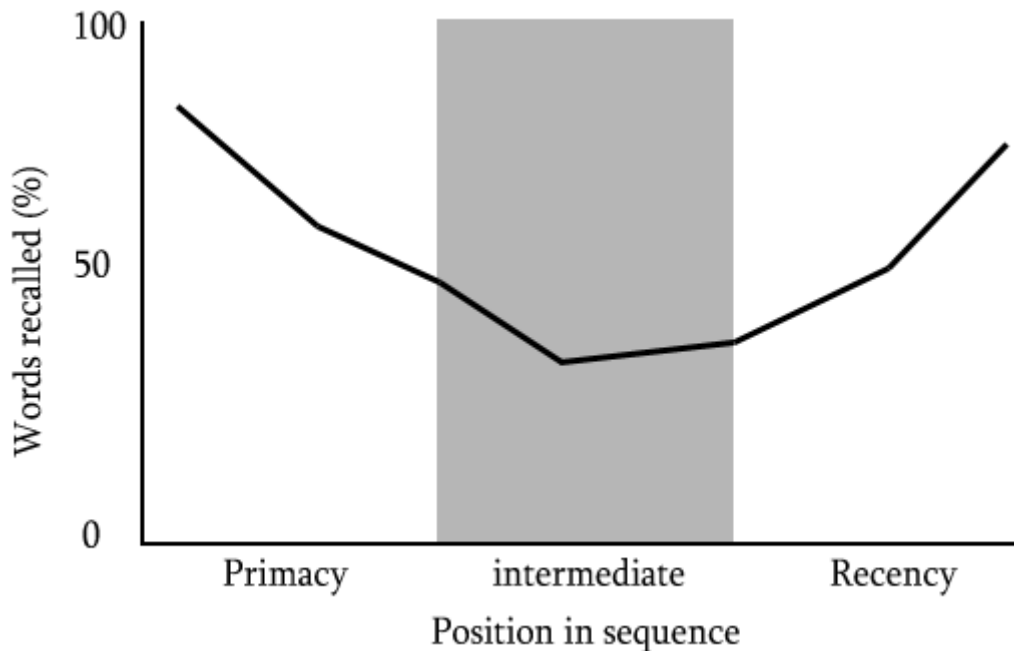
Graph showing the serial position effect. The vertical axis shows the percentage of words recalled; the horizontal axis shows their position in the sequence.

In memory, in contrast, early accounts of primacy and recency focused on proactive and retroactive interference respectively (e.g. Foucault, 1928; cited in Murdock, 2001). Primacy effects have been (non-exhaustively) attributed to additional rehearsal of early-list items (Rundus, 1971), more distributed or more recent rehearsal of early-list items (Modigliani & Hedges, 1987; Tan & Ward, 2000), various measures of distinctiveness (Johnson, 1991; Murdock, 1960), anchoring effects (Feigenbaum & Simon, 1962; Glanzer & Dolinsky, 1965), inhibitory processes (Hull, 1935; Lepley, 1934) or reducing encoding throughout a list (Brown et al., 2000; Farrell & Lewandowsky, 2002; Lewandowsky & Murdock, 1989; Page & Norris, 1998). Recency effects over short time-scales have been (non-exhaustively) attributed to primary memory (Waugh & Norman, 1965) or to response selection processes (Farrell & Lewandowsky, 2002; Lewandowsky, 1999; Lewandowsky & Farrell, 2000; Lewandowsky & Murdock, 1989), whereas long-term and short-term recency effects have been attributed to various temporal discriminability mechanisms (Baddeley, 1976; Baddeley & Hitch, 1993; Bjork& Whitten, 1974; Crowder, 1976; Glenberg et al., 1983; Glenberg & Swanson, 1986; Neath, 1993a,b; Tan & Ward, 2000). Thus in recent years the goal of developing a unified approach to serial position effects in memory and absolute identification has effectively been

abandoned. Even within traditional serial and free recall paradigms, different explanations for primacy effects, short-term recency effects, and long-term recency effects survive in parallel.

Scale Invariant Effects in Memory and Perception

In addition to focusing on serial position effects in a range of tasks, the present paper aims to model data (from both absolute identification and memory paradigms) that hold over many different scales. What reason is there to believe that such an account is either necessary or possible? We noted earlier that scale-invariant explanations are often preferred in science. Here we describe scale-invariant phenomena in memory and absolute identification, then emphasize the importance of scale-invariance for psychology more generally (Chater & Brown, 1999). The concept of scale-invariance is not always used consistently, and has technical usages that we do not require here, but in intuitive terms can be taken to characterize any system whose characteristics are similar independently of the scale of measurement (for accessible introductions see Bak, 1997; Flake, 1998; Mandelbrot, 1982; Schroeder, 1991). Scale-invariant physical structures are said to be statistically self-similar, in that small parts of an object are on average similar to the whole. The central point here is that the statistical structure of a data set (such as the degree of wiggliness of a river) provides no information as to the scale that is being examined. Natural-world examples are provided by structures such as clouds, or the coast of Norway - in both cases an observer would be unable (in the absence of any other cues) to distinguish between a small nearby structure and a larger distant structure. In human memory, an analogous example detailed below is the ratio-dependent nature of recency effects such that examination of a recency effect will not in itself be informative regarding the scale of the effect (i.e., whether it is a long-term or a short-term recency effect).

Scale Invariance in Memory

Evidence for scale-independence in human memory may take a number of different forms. Particularly relevant data come from results that implicate similar processes operating over different timescales; such results motivate models like the one presented here in which relative amounts of time, not absolute amounts of time, influence retrievability.

First of all, evidence for power-law forgetting would be consistent with scale- independence in memory performance over time. If the forgetting curve does follow a powerlaw, then the probability of recall will depend on Ta where T is the amount of time since an episode was learned and a is a constant. To the extent that the time course of human memory loss does more or less closely follow a power-law, as has been suggested by a number of researchers (see e.g. Anderson & Schooler, 1991; Rubin & Wenzel, 1996; Wixted and Ebbesen; 1991, 1997), forgetting may be seen as scale- independent.[1] Many researchers have however claimed that forgetting curves are not best described by a power law (for recent examples see Rubin, Hinton, & Wenzel, 1999; Wickens, 1999; cf. also Myung, Kim, & Pitt, 2000); we address this issue below.

Further evidence consistent with scale-independence has come from the study of recency effects. Although recency effects disappear after a filled retention interval (Glanzer & Cunitz, 1966; Postman & Phillips, 1965), the effect reappears if the spacing between presented items is increased (Bjork & Whitten, 1974) and is seen when retrieval from LTM is required (Baddeley & Hitch, 1977; Pinto & Baddeley, 1991; Sehulster, 1989). More generally, the size of the recency effect appears to depend on the log of the ratio between the IPI (inter-presentation interval between the items) and the RI (retention interval) (Glenberg et al. 1980; Nairne, Neath, Serra, & Byun, 1997), at least to a remarkably close degree (see also Baddeley, 1976; Bjork & Whitten, 1974). The relation holds even when the IPI to RI ratio varies over many orders of magnitude, from milliseconds to weeks. Empirically, the ratio rule means that it would be impossible for an observer to judge from the size of a recency effect whether that recency effect arose from recall of a list of items presented 10 s apart and followed by a 50 s retention interval, or a list of items presented 1 s apart and followed by a 5 s retention interval.

A third line of evidence for scale-independence in memory is the fact that the proportion of errors produced for each serial position during serial learning remains constant even in the face of considerable variations in degree of learning, IPI, time between trials, familiarity or meaningfulness of the material to be remembered, or individual differences in the learners (Braun & Heymann, 1958; McCrary & Hunter, 1953);

this is the **Hunter-McCrary Law**. Here the shape of the error distribution in serial learning provides the observer with no evidence about the absolute level of performance. We present representative data below.

Fourth, serial position effects in rather different tasks, such as order reconstruction, remain qualitatively (and sometimes quantitatively) unchanged at different time scales (e.g. for the dimensions of position-within-list and list-within-trial: Nairne,1991; data below). Similar qualitative scale-invariant features are evident in the data from grouping experiments, where items at the beginning and end of each group are better recalled, echoing the primacy and recency effects for the list as a whole (Frankish, 1985, 1989; Ryan, 1969a, 1969b; Hitch, Burgess, Towse, & Culpin, 1996).

Fifth, the pattern of transposition errors in order reconstruction tasks remains remarkably constant across timescales varying over many orders of magnitude, from milliseconds to weeks (see Huttenlocher, Hedges, & Prohaska, 1992; Nairne, 1991, 1992; Neath, 1998; see Brown et al., 2000, for a summary). A sixth line of evidence comes from the study of memory for relative recency. Underwood (1977) asked participants to recall the dates of events that occurred between four months and seven and half years in the past, and found that the greater the time separating two events, the less likely those events were to be recalled in the wrong relative order. Hacker (1980) found similar effects for items separated by just a few seconds. Seventhly, there is evidence for invariance over the time-frame of recall. Maylor, Chater, and Brown (2001) asked participants to recall events from the past day, week, or year. The cumulative response probabilities were indistinguishable across the three conditions. Finally, scale-invariance is also evident in recall-order effects in free recall (Howard & Kahana, 1999).

In summary, there is considerable evidence that many important properties of memory are scale-invariant, in that similar effects are evident at many different timescales (see also Melton, 1963; Nairne, 1992; 1996). An emphasis on scale-invariance resonates with recent models of scale-invariant processes in animal learning; Gallistel and Gibbon (2000) describe a model of scale-invariant conditioning that is closely related to the model of human memory proposed here. Models of timing also emphasize scale-invariance (e.g. Killeen & Taylor, 2000; Wearden, 1994).

Scale Invariance in Absolute Identification

Scale-invariance is also observed in absolute identification. In a typical absolute identification task, participants are exposed to a set of stimuli arrayed along some dimension such as pitch, weight, area, or amplitude (e.g. nine tones of different frequencies, in the experiments we report below). A label is associated with each stimulus. The labels may be numbers (e.g. 1 through 9), with the number for each item corresponding to the item's ordinal position on the continuum, or may be arbitrary (e.g. the names of different colors may be associated with the different stimuli, in which case the task essentially becomes more akin to paired associate learning). Participants are then exposed to individual stimuli in random order and required to identify them with the correct label. Feedback regarding the correct response is normally given after each trial.

We present data from several absolute identification studies below; here the central (and counterintuitive) point is that identification performance for pairwise- discriminable items is often almost unaffected when the spacing of items along the perceptual scale is increased by a constant factor (see, e.g., Alluisi & Sidorsky, 1958; Eriksen & Hake, 1955; Garner, 1962; Miller, 1956; Pollack, 1952; Shiffrin & Nosofsky, 1994). This provides a clear example of scale-invariance. Further scale-invariance is evident in the serial position effects obtained in absolute identification experiments, for such curves have essentially the same form no matter what the range of the relevant perceptual dimension that is used (Neath, Brown, McCormack, Chater, & Freeman, 2006).

Scale Invariance and Psychological Explanation

Scale-invariant phenomena appear to call for scale-invariant models. Is a bias towards scale-invariant models inconsistent with memory data? In the study of human memory, it has been assumed that the data do not permit scale-invariant explanations. Distinctions between the memory retrieval principles that operate over short and long time scales have a long and distinguished history (Atkinson & Shiffrin, 1968;

James, 1890/1950; Waugh & Norman, 1965) and are still widely accepted (see, e.g., Baddeley, Gathercole, & Papagno, 1998; Cowan, Wood, & Borne, 1994; Gathercole, 1999; Izawa, 1999). In particular, time-based trace decay is assumed to be an important factor underpinning forgetting over short timescales, with interference being assumed to predominate over longer timescales. Our exploration of scale-invariant retrieval principles is not of course intended to rule out the assumption of a separate STM (although we do suggest that much existing evidence can be interpreted within a scale-invariant framework). In focusing on aspects of memory that appear similar over widely varying time scales we attempt a new perspective on an old tradition of suggesting that the same interference-based principles apply over both short and long time scales. We call this the scale-invariant memory assumption. The present proposals regarding scale-invariance in human memory and absolute identification particularize a more general claim that scale-invariance in cognitive and perceptual function may reflect scale-invariance in the structure of the environment (Chater & Brown, 1999; cf. Anderson, 1990; Shepard, 1987a).

Time, Identification, and Categorization

Our concern is with temporal scale invariance in the case of memory, but with invariance over stimulus range in the case of absolute identification. A common perspective is sought - do the same principles govern the discriminability (and hence retrievability) of items in memory as govern the discriminability of stimuli from one-another in absolute identification paradigms? More specifically, do serial position effects arise for the same reasons in memory and in identification (Murdock, 1960)? The relationship between absolute identification and serial and free recall remains unclear, and in recent years different theoretical and modeling approaches have been taken in explaining these different paradigms. Current models of absolute identification, categorization, and recognition performance account for a range of empirical data to a high level of precision (e.g. Ashby, 1992; Ashby & Perrin, 1988; Erickson & Kruschke, 1998; Estes, 1994; Kruschke, 1992; Kruschke & Johansen, 1999; Lamberts, 1995; Nosofsky, 1986; Nosofsky & Palmeri, 1997). In many respects3 such models seem more advanced than current models of serial and free recall. However the insights embodied in models of identification and classification have not generally been applied to traditional serial and free recall memory paradigms. Here we suggest that this is partly because multidimensional scaling models of categorization have not included time as an important dimension underpinning memory retrieval. Although models can allow for the differential availability in memory of exemplars 4, the relation between temporal factors and memory/exemplar availability has not been widely explored. However in recent models of memory, and of serial recall in particular, time has been accorded a central role (e.g. Altmann & John, 1999; Anderson & Matessa, 1997; Anderson, Bothell, Lebiere, & Matessa, 1998; Brown et al., 2000; Burgess & Hitch, 1996; 1999; Houghton, 1990; Neath, 1993a, 1993b; see also Hintzman & Block, 1971; Hintzman, Block, & Summers, 1973) and in more general terms it has long been argued that time is an important dimension underpinning memory organization and retrieval (e.g. Gallistel, 1990; see Brown & Chater, 2001, for a recent review; Friedman, 2001, for an alternative perspective). Here we apply simple principles of the type previously explored mainly in categorization and identification models to serial recall and free recall paradigms via the addition of a temporal distance dimension into such models.

(b) Evaluate 'duplicity theory' of vision. Cite experimental evidence in support of this theory.

TOPIC: *Sensation, Attention and Perception*
SUBTOPIC: *Biological Factors in Perception*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

Duplicity theory

Central to any discussion of dark adaptation is the duplex nature of the anatomical organisation of the human retina. The duplicity theory of vision is based on the fact that the human retina contains two different photoreceptors, rods and cones, which function independently of each other, operating under different conditions, and giving rise to qualitatively different perceptions. The average human retina

contains 4.6 million cones and around 92 million rods. In the periphery, the rods dominate; more centrally the relative number of cones increases and in the fovea, the visual centre of the eye, there are virtually only cones. A graphical representation of the spatial density of rod and cone photoreceptors within the human retina is shown in Figure 1. Cones facilitate day vision, colour vision and provide more detailed spatial resolution. Rods facilitate night and peripheral vision, but have relatively poor ability to discern detail or motion and are insensitive to colour.
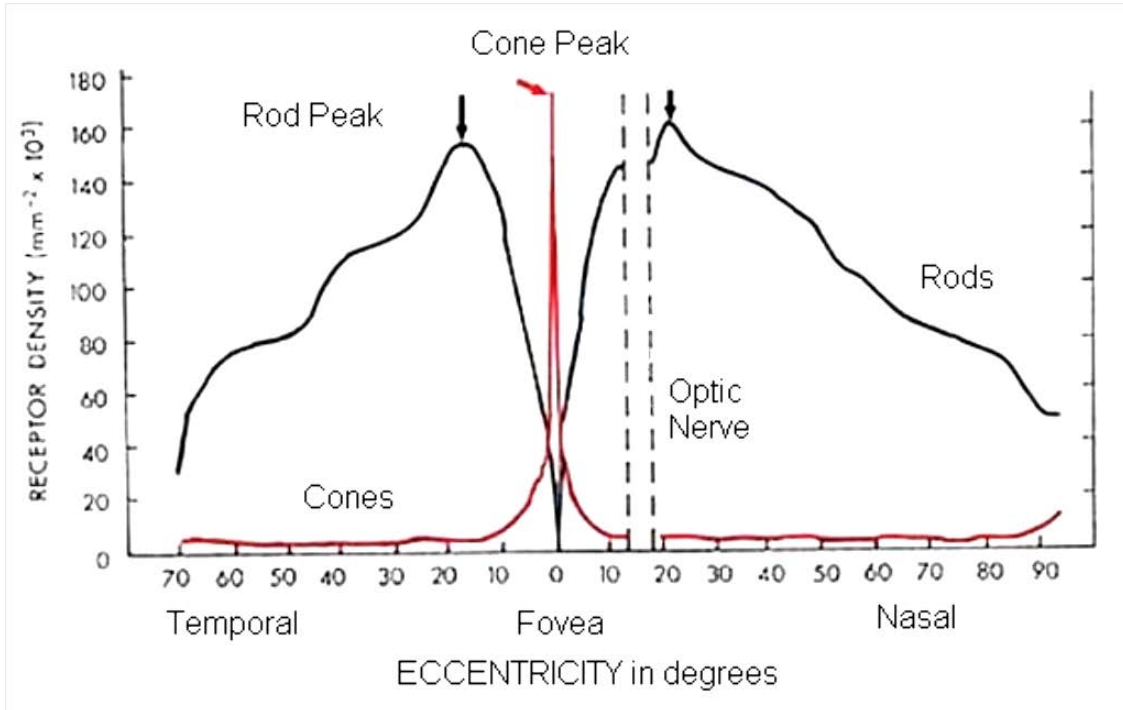

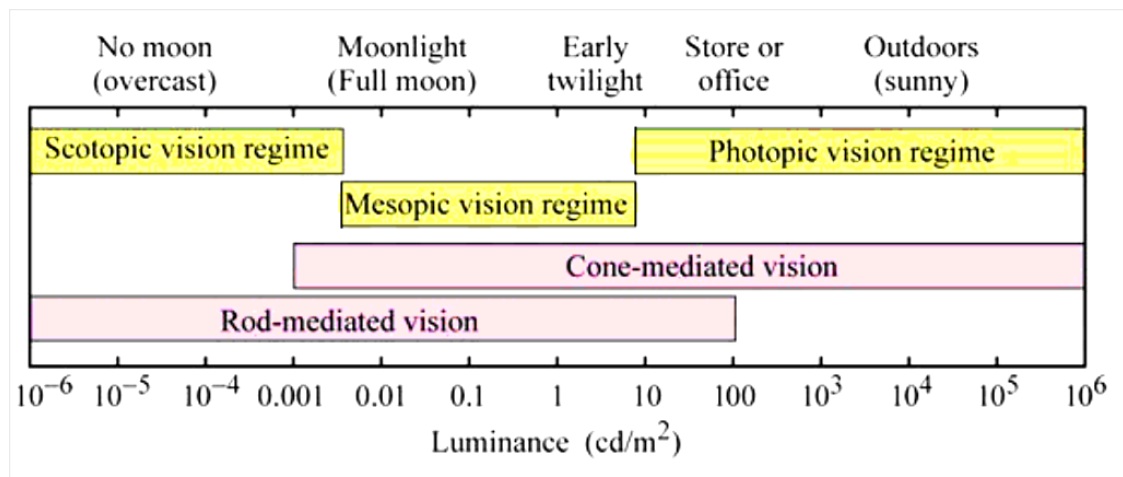**Figure 1** Distribution of rods and cones in the human eye. From G. Osterberg

Figure 2 Range of conditions that can be accommodated by the human eye

Rods and cones function exclusively within certain illumination ranges. Under reduced illumination, ambient lighting below 0.03 cd/m2 (candela per square metre), rods cells are functional and exclusively mediate vision; this is referred to as scotopic vision (often, incorrectly, termed 'night vision'). As the light increases, rod function starts to decrease and the cones begin to operate. Once luminance increases sufficiently, so that only cones are functioning, the visual system is operating in the photopic range. The transition between photopic (cones) and scotopic (rods) vision is gradual; both the rods and cones as well as the adaptive mechanism of each system interact in the mesopic range. Figure 2 illustrates the range of conditions that can be accommodated by the human eye.

Classic Dark Adaptation Curve (DAC)

A clear illustration of the duplex nature of the visual system is shown in the classic Dark Adaptation Curve (DAC) reported by Hecht and Mandlebaum, who observed a two stage decrease in visual threshold with time after initial pre-adaptation to a very bright light source. The methodology used to obtain a DAC typically involves pre-exposing participants to a very bright light source, so that they become fully light adapted, after which the light is turned off (an instant called time zero), and a test flash is presented at various intervals to designated retinal locations. The minimum luminance (threshold) required to detect the test flash is determined by adjusting the intensity of a target stimulus (usually a small test flash of less than 1 second) until participants can just see it. Threshold is measured using the method of ascending limits (from non-seeing to seeing). The first observation is made within the first 15 to 30 seconds after the termination of the bleaching light, and then as many readings as possible are taken until a stable threshold is obtained. It is widely accepted among the scientific community and has been consistently replicated using a variety of methods and populations.
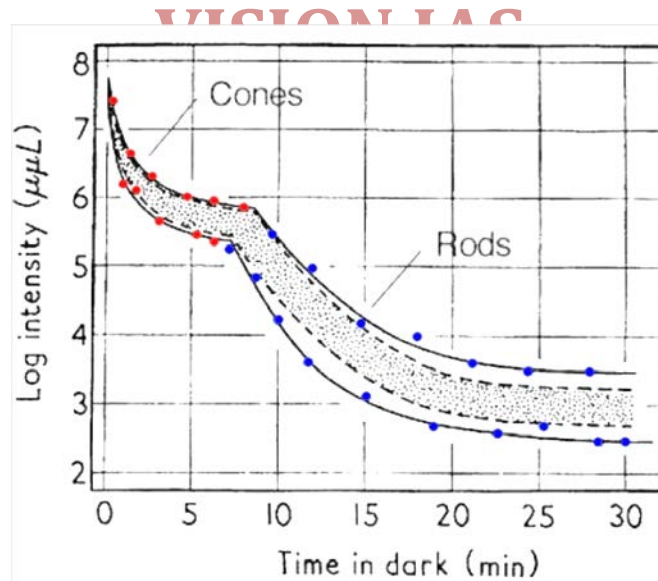
Figure 3 Dark adaptation curve. Hecht and Mandelbaum (1939). From Pirenne M. H., Dark Adaptation and Night Vision. Chapter 5. In: Davidson, H. (ed), The Eye, vol 2. London, Academic Press, 1962

Figure 3 illustrates the time course of dark adaptation; the shaded area represents the data for 80% of Hecht and Mandlebaum's participants (N = 110 university staff, students and volunteers). Time is plotted on the horizontal axis on an ordinary linear scale. Light intensity, however, is plotted on the vertical axis on a power or logarithmic (base 10) scale. On such a scale 1 means 10 units, 2 means 100 units, 3 means 1,000 units, etc. (see Box 1 for a definition of the units used in photometry).

The first phase of the curve reflects the fact that, after exposure to a bright light (providing a strong bleach), the rods are relatively insensitive and it is the cone system which is the more sensitive of the two. As a consequence, it is the cone system that is initially the active system in determining the threshold for the detection of the light. Depending on the strength of the bleach, the first phase may take around 5 to 10 minutes to complete, approaching a minimum threshold in the mesopic range. Phase one of the curve was originally conceptualised as a hypothetical limit obtained if cone activity could be measured independently of rod activity, and making the basic assumption that the potential cone threshold reaches a stable plateau level after a few minutes of dark adaptation.

More recent research has supported this position, for example, Stabell and Stabell report that the DAC of the rod-free fovea reached a stable plateau after approximately 8 minutes. After around 5-10 minutes in the dark, the rod-cone break occurs, i.e. the sensitivity of the rod pathway improves to the point where the rod system becomes the more sensitive of the two. This is reflected by the second part of the DAC (often referred to as the S2 component), wherein the rods require around 30 to 45 minutes to effectively reach their minimum (absolute threshold) of around 10-5 cd/m2. Further evidence to support the rod-cone break is again provided by Stabell and Stabell who demonstrate changes in the perception of colour during dark

34

adaptation. During the cone-plateau period, the colour mixture functions remained invariant. However, at the point where the rod-cone break was observed, the perception of colour changed. When the rod mechanism takes over, coloured test spots appeared colourless and, as such, it can be concluded that the second component of the DAC is due to pure rod activity as only the cone pathway provides the sensation of colour.

Changes in stimulus dimensions = changes in dark adaptation

In is important to note that the shape and time course of a dark adaptation curve is dictated primarily by the duplex anatomic structure of the retina and the extent to which rods, cones, or both are influenced by the adapting light and by the test flash. Fundamental differences exist between experimental laboratory studies that measure light threshold of the dark adapted eye and substantial variations can occur with small changes in stimulus dimensions. In any determination of a DAC, the actual value of threshold intensity will depend upon the size, duration, wavelength etc of the test flash.

Particularly relevant in the examination of the lookout task is the intensity of the pre-adapting light source. It is highly unlikely that the bridge crew will, immediately prior to commencing night time lookout duties, be exposed to a light source comparable to those used in the laboratory study of dark adaptation. They are far more likely to adapt from much lower levels of light and, as seen in Figure 4, decreasing the light levels prior to adaptation reduces the initial threshold of detection and decreases the time taken to reach both the rod-cone break, and the minimum detection threshold. Furthermore, if the adaptation light source was 3,000 Td or less, only the rods are sensitive enough to detect the test light, as evidenced by the absence of a rod-cone break, and significant improvements in minimum threshold occur during the first 10 minutes, and the final steady-state level is almost reached after about 15 minutes.
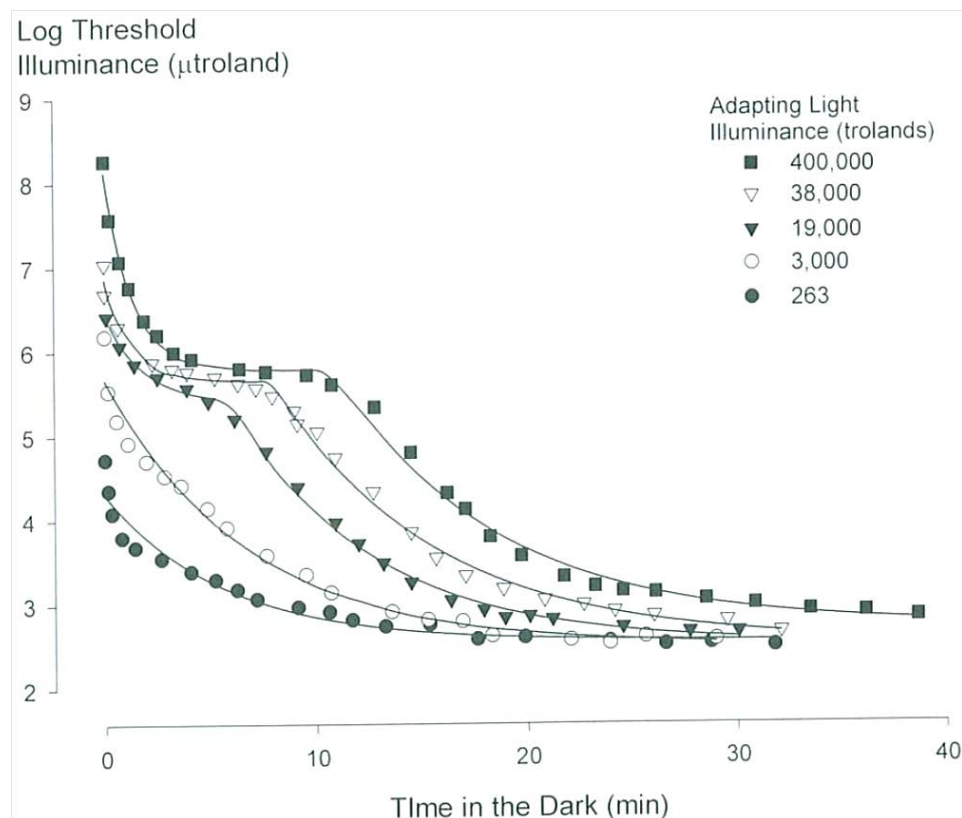
Figure 4 Dark adaptation curves following different levels of pre-adapting luminance (In 5; Modified from 31)

THE ROLE OF RHODOPSIN IN DARK ADAPTATION

Much of dark adaptation can be explained by a simple photochemical process in which the photopigment rhodopsin (originally referred to as visual purple) is bleached away during exposure to a bright light, and is re-synthesised in the dark. Rhodopsin is a chemical combination of vitamin A (in the form of 11-*cis* retinal) and an opsin, a group of light sensitive protein receptors found in the retina. Rushton et al used retinal densitometry, which allows the level of visual pigment to be measured in the living eye during bleaching and regeneration, to plot the time course of the regeneration of human rhodopsin; they found that it takes around 40 minutes following a full bleach to completely regenerate. Similarly, the level of photopigment in the cones (chromopsins) return to normal in around 8 minutes. Both these times correspond closely to the times detailed in the classic dark adaptation curve shown in Figure 3, which strongly suggests that there is a direct relationship between threshold and the concentration of photopigment. Furthermore, Rushton illustrates that despite very different extents of bleaching and subsequent periods in the dark, the rod-cone break appears to occur when around 92% of rhodopsin has been re-synthesised. This is summarised in Figure 5, which shows that after total bleach (O), the moment when the rods are first sensitive enough to respond to a green flash occurs after around 15 minutes, corresponding to the 92% threshold. Subsequent repetitions with weaker pre-adapting light (O = 25%, ● = 35%) produced shorter adaptation times. However, when adjusted so that the first point of the curve (zero time) falls onto the exponential curve, not only do all subsequent points lie upon the same curve, but the point when participants first respond to the test stimulus is the same, i.e. 92% rhodopsin. Additional evidence of the link between the recovery of visual sensitivity and the regeneration of rhodopsin can be found in Lamb and Pugh.
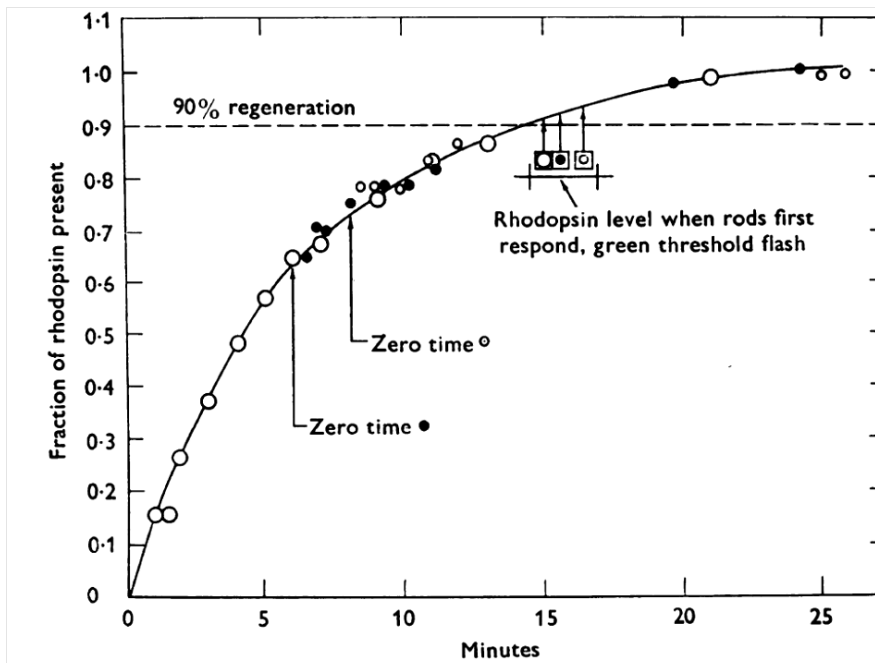

Figure 5 Regeneration of rhodopsin during dark adaptation. The rod-cone break for each condition is marked with an arrow

NEURAL ADAPTATION

Adaptation in the eye not related to the visual pigment level is usually referred to as neural adaptation. There are aspects of visual function that change during dark adaptation that cannot readily be accounted for in terms of photochemical changes, but reflect changes in the response of the photoreceptors and the cells that receive input from the photoreceptors. Rapid neural adaptation has been observed during the initial stages of dark adaptation when measured from the increment or instantaneous threshold level (7;36;37). For example, Baker examined the threshold for detecting the test flash within 2 seconds after time zero and found that there is a rapid decrease in threshold within 0.3 seconds after the adapting flash was turned off (57, 1,800 and 57,000 trolands for 20 milliseconds centred on the fovea).

Lythgoe first drew attention to the fact that lights which bleach only a small fragment of visual purple (rhodopsin) could still cause very big changes in threshold and suggested that the change is in part due to the alteration of nerve organisation. It was hypothesised that the decreasing ability to perform fine visual judgements under scotopic conditions may be due to synaptic rearrangement. During complete light adaptation, each visual element is served by separate nerve fibres, which make the individual capable of performing fine judgements, while during dark adaptation each fibre serves several elements by a spread of its synaptic connections, which reduces discrimination of fine detail but improves retinal sensitivity to light.

It is generally agreed that a rod can be excited by the catch of a single quantum. That is, 1 quantum of light absorbed can excite a dark-adapted rod in the human eye and a fully dark-adapted human subject is capable of detecting as few as 5–10 photon absorptions occurring within a short time interval anywhere over a 'pool' of around 10,000 rods (41;42). Therefore, adapting lights 70,000,000 times above the visual threshold should not bleach away more than 2 per cent of the visual pigment in a rod over the course of a 5 second exposure, raising the visual threshold by less than 0.5 of a log unit. Yet evidence suggests that adapting lights of this luminance raise the human visual threshold 2 to 3 log units when measured at the beginning of dark adaptation. Rushton hypothesised that rod signals converge upon a 'summation pool' (illustrated in Figure 6), which responds to the total flux of signals, independent of the particular rods excited. Whenever a rod absorbs a quantum it generates one signal that travels to the 'summation pool'. If a critical number of signals (N) arrive at the pool within a short time interval, a message will be relayed to the brain and the light will be seen. The total flux of these messages modifies the sensitivity of the pool and determines the sizes of flux increment needed for detection.

Further evidence of neural changes during dark adaptation can be found in a consideration of a number of studies that suggest adaptive effects can spread laterally across the retina. Illumination in one area of the retina has been shown to elevate thresholds for test stimuli falling on areas that have not be directly exposed to light. For example, Lipetz projected a small conditioning spot onto the excised retina of a bullfrog for several minutes and then measured the ganglion cell thresholds for a test stimulus at the previously illuminated area and at an unilluminated area. The unilluminated area showed an increase in threshold, which was greater than could be explained by scattered light.
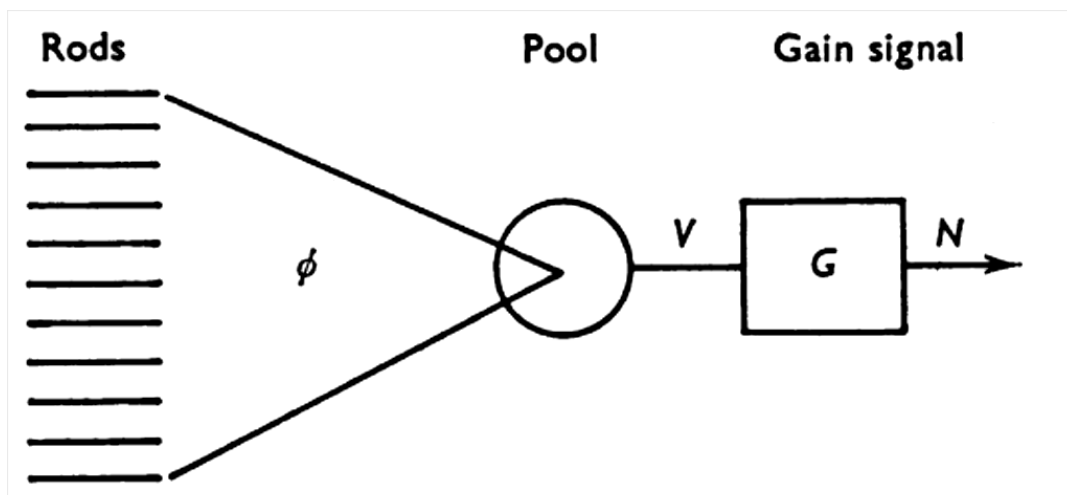
Figure 6 The effect of a flash of light (Ø), the pooled rod response (V) is modified by a variable gain box (G) to produce the output (N)

Rushton and Westheimer were able to reach similar conclusions in experiments on man. The aim of their investigation was to find out whether the rod threshold rises after bleaching rhodopsin because (a) the rods need more light to generate a signal, or (b) because the summation pool needs more signals to activate the optic nerve. To achieve this, a grating of black and transparent stripes of equal size was focused upon the retina so that the retina was 'bleached' and 'spared' in parallel strips and was compared to a second similar region of the retina bleached using a neutral density of 0.4, designed to transmit the same total light as the grating. Dark adaptation in the two regions was measured with the test flashes reversed, i.e. the uniformly bleached area was tested with the grating interposed, and the stripe-bleached area with the

density filter interposed. Results show that the loss of sensitivity after bleach is not confined to the area exposed to the adapting light source, and the two dark adaptation curves were practically identical. The results did not support mechanism (a), that the rods need more light to generate a signal, as the strip of retina that lay under the dark bar of the grating image received less illumination than the rest of the test spot, yet threshold levels were not detectably lower. This suggests that the threshold of a rod may be raised by bleaching the rhodopsin of neighbouring rods, which would be expected if the factor that changes during dark adaptation is the threshold of the pooled rod signals, i.e. mechanism (b). In sum, these experiments appear to establish that, as Rushton hypothesised, changes in retinal sensitivity are controlled by signals pooled from many receptors.

It is now known that the synaptic ratio between the photoreceptors and bipolar cells, which transmit signals from the photoreceptors to the retinal ganglion cells that form the optic nerve, plays an important part in both night vision and visual acuity. The centrally located cones synapse 1:1 with the bipolar cells, and are ultimately represented individually in the cortex. As one moves more peripherally, a larger number of primary receptor cells synapse with each bipolar cell. Scotopic vision is mediated by two distinct pathways from rods to ganglion cells that operate at different light intensities. At higher scotopic levels, rod signals are electrically coupled via gap junctions to cones, and then follow the conventional photopic pathway via cone bipolar cells to ganglion cells. But at low scotopic intensities a separate pathway comes into play, capable of transmitting single photon responses. The rods synapse onto rod bipolar cells, which are functionally similar to cone bipolar cells. Under these conditions, if the rods are slightly stimulated, the summation of several low level stimuli may be enough to excite a bipolar cell, thus sending a 'light' signal to the brain.

(c) Examine the impact of cultural factors on socialization process among children.

TOPIC: *Development of Human Behavior*
SUBTOPIC: *Influence of Cultural Factors in Socialization*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

Humans are social animals, and much of our everyday lives involves interactions with and influences of others. Especially in North America, we struggle with the tension between being unique, separate individuals ("Every man for himself ") and being fundamentally connected to one another ("No man is an island").

Culture exerts considerable influence in the social arena. How we interact with others, perceive others, and work with others—all are influenced by the culture in which we live. We have all learned particular ways of behaving, perceiving, and working based on our own cultural upbringing and milieu. We may believe our way is the way people of all cultures should behave and interact, but what is true for us is often not true for people of other cultural groups.

Culture and In-group/Out-group Relationships

Culture influences social behavior in many different ways. We all live with others, forming attachments, bonds, and relationships. We are close to some people and distant from others. We make friends, acquaintances, and even enemies. Some of the people we see every day we know well, yet other people we see every day we don't know at all. Strangers, family members, friends, coworkers, acquaintances—the list of people in our everyday world is long.

One way social scientists have learned to understand our relationships with different people is by classifying them into categories that approximate the psychological categories we create. Especially important to understanding self–other relationships and pertinent to understanding cultural differences in social behavior is the category of in-groups versus out-groups.

In-groups and Out-groups

The in-groups–out-groups classification is one of the oldest and best studied social classifications in social psychology and sociology (see Brewer & Kramer, 1985; Messick & Mackie, 1989; and Tajfel, 1982, for reviews and more complete descriptions of this distinction). Most of us intuitively know the difference between in-groups and out-groups. In-group relationships are relationships characterized by some degree of familiarity, intimacy, and trust. We feel close to the people around us we consider in our in-group. Self–in-group relationships develop through bonds that tie the in-group together through common friendship or relationships or goals.

Out-group relationships are just the opposite. Out-group relationships lack the familiarity, intimacy, and trust characteristic of relationships with in-group others. In-group relationships may be associated with feelings of closeness, but out-group relationships may lack such feelings altogether and may even involve negative feelings of hostility, aggression, aloofness, or superiority. A bond exists that binds in-group relationships together, but no such bond exists for relationships with people on the outside. These people simply exist and are barely in our consciousness. They do not have any special relationship with us.

The in-group –out-group distinction is dichotomous, allowing us to characterize or classify everyone in our world into one of these two categories. But social scientists know the world is not that simple. Our social relationships cannot be neatly classified into two categories. There are differing degrees of intimacy, familiarity, and closeness even within one category. Classification schemes like in-groups–out-groups are simply aids that help us understand our behavior with others while acknowledging that greater complexity exists in those relationships.

Much of socialization and enculturation—the time of growing and learning about the rules and standards of our society and culture—is spent learning which people constitute our in-group and out-group . From birth (and arguably before), we are busy building relationships with the people around us. As we go to school, make friends, find jobs, fall in love, and generally go through life, we develop relationships with many different people. Explicitly or implicitly, we categorize those relationships in our own minds according to the dimensions that define our in-groups and out-groups.

The in-group –out-group distinction is applicable to all cultures and societies of the world. People of all cultures must learn to differentiate among the people they have relationships with. However, people of different cultures differ in exactly how these relationships develop, and with whom. The people we generally consider to belong to our in-group may not be the same people that members of another culture consider to be in their in-group . The same is true for out-groups. And even when the same people can be classified as in-group or out-group across cultures, the particular shapes, forms, and meanings of those relationships may be entirely different. Recognizing these similarities and differences forms the basis for understanding how culture can influence in-group and out-group relationships and guide our social behaviors.

Cultural Differences in In-group/Out-group Relationships

Structure and format of in-group /out-group relationships. We have already touched on how people of different cultures can differ in their self–in-group and –out-group relationships. Our own observations suggest that people of different cultures may not consider the same types of people and relationships when defining in-groups and out-groups. Just because we consider a certain type of person (a friend at school, a work colleague) an in-group (or out-group) member, we cannot assume that people from another culture will interpret and act on those relationships in exactly the same way.

Cultures differ in the formation and structure of self–in-group and self–out-group relationships in other ways as well. In North American culture, in-group and out-group membership is stable, no matter what we are talking about, to whom we are talking, or where we are talking. Our friends are our friends no matter what. But to someone in another culture, some people may constitute an in-group in one circumstance or situation, but the same people may constitute an out-group in another. It is not uncommon for businesspeople in many Asian cultures, for example, to consider each other out-groups and competitors

when talking about domestic business issues. But when the discussion turns to international business competition, those same out-group competitors may band together to form an in-group. This type of switching of in-groups and out-groups is not limited to Asian or collectivistic cultures; it is present in many, if not all, cultures. When former President Bush visited Japan in 1991 with the chief executive officers of many different American companies, they all represented in-group "Americans," even though those companies and officers would consider each other out-group rivals in relation to domestic issues. As with many cultural differences, cultures differ in terms of degree but not necessarily presence or absence of this switching phenomenon.

Cross-cultural research has amply demonstrated that people of different cultures perceive relationships differently. Forgas and Bond (1985), for example, asked participants in Hong Kong and Australia about 27 social episodes related to university student life. They asked participants to group the episodes into categories based on their perceived similarity, to label their groupings, and to identify subgroups of the most similar episodes within those groups. They then subjected the data to a statistical technique known as multidimensional scaling, which identifies the kinds of dimensions that best underlie the groupings made by each group of participants. Four dimensions best described how the Hong Kong Chinese participants cognitively represented the episodes: (1) power distance, (2) task versus social orientation, (3) evaluation, and (4) involvement. For the Australian sample, however, a different set of four dimensions emerged: (1) competitiveness, (2) social versus task orientation, (3) involvement, and (4) self-confidence. These differences are clearly related to differences in the cultural values of the two societies. The Hong Kong Chinese cultural values of strong communal feelings, collectivism, and social usefulness, together with an acceptance of authority, influence their perceptions of the social events. For the Australians, competitiveness, self-confidence, freedom, and hedonistically based evaluations dominate their perceptions. These data demonstrate clearly how the same social episodes can be perceived very differently in different cultures.

Another recent study also demonstrated interesting psychological differences in how people perceive their in-groups. In this study (Harrison, Stewart, Myambo, & Teveraishe, 1995), adolescents in Zimbabwe and the United States completed a 33-item test that measured six aspects of social relationships: (1) reliable alliance, (2) enhancement or worth, (3) affection, (4) instrumental help and guidance, (5) companionship and social integration, and (6) intimacy. The researchers added three other dimensions: (7) conflict, (8) satisfaction, and (9) discipline. The participants completed this test about their relationship with six target individuals: mother, father, favorite relative, teacher, best friend, and favorite sibling. The results indicated that the Zimbabweans perceived social support as being provided by a variety of persons in their social network. The Americans, however, perceived social support as being provided primarily by parents and best friends. Also, the Zimbabweans perceived their social support network as providing them with intimacy; the Americans, on the other hand, perceived their social support network as basically providing them with affection. Again, these differences highlight the different ways in which people of different cultures can perceive in-groups and out-groups, and the different psychological meanings attributed to them. These differences are related to differences in cultural values: Zimbabwean culture places a higher value on relationships, whereas American culture places a higher value on individuality and uniqueness.

The meaning of in-group/out-group relationships. Some scientists have done a considerable amount of work on cultural differences in self–in-group and –out-group relationships. Triandis and his colleagues (1988) have done an especially nice job of elucidating how self–in-group and self–out-group relationships differ across cultures by using the cultural dimension of individualism versus collectivism to understand cultural differences in social behavior.

As we have seen throughout this book, individualism–collectivism (IC) is one of the most important social psychological dimensions of culture. Many writers across the social science disciplines have used this dimension to understand differences in social behaviors across the cultures they have studied (for example, Hofstede, 1980, 1983; Kluckholn & Strodtbeck, 1961; Mead, 1961; Triandis, 1972). IC refers to the degree to which a culture promotes individual needs, wishes, desires, and values over group and collective ones. Individualistic cultures encourage their members to become unique individuals; hierarchical power and status differences are minimized, and equality is emphasized. Collectivistic cultures stress the needs of a group; individuals are identified more through their group affiliation than by individual position or

40

attributes. Hierarchical differences and vertical relationships are emphasized, and role, status, and appropriate behaviors are more clearly defined by position. Self–in-group and self–out-group relationships differ in individualistic and collectivistic cultures, and these differences in the meaning of in-group and out-group relationships produce differences in the types of behaviors people engage in when interacting with others. In individualistic cultures, such as the United States, people often belong to multiple in-groups. Many Americans belong to several in-groups—music groups, sports groups, church groups, social groups, and so forth. Children may belong to football teams during football season, basketball teams during basketball season, and baseball teams during baseball season. They may take swimming lessons, take piano or violin lessons, belong to Boy or Girl Scouts, and generally just be the busiest people around. In contrast, members of collectivistic cultures, including many Asian and South American cultures, belong to fewer in-groups. They do not belong to all the different sports, music, and social groups that people in individualistic cultures do.

This difference between individualistic and collectivistic cultures in in-group membership has important consequences for the degree of commitment people have to different groups. In general, in exchange for belonging to fewer groups, people in collectivistic cultures have greater commitments to the groups to which they belong. They also identify more with the groups to which they belong; that is, the groups themselves become an integral part of each individual's self-concept and identity. This makes sense because, by definition, collectivistic cultures depend on groups to a much greater degree, and subjugating personal goals in favor of collective goals is a necessity.

Members of individualistic cultures do not necessarily collapse their sense of self-identity and self-concept into the groups to which they belong. They have fewer commitments to their in-groups and move much more easily from in-group to in-group. Groups take on special importance in collectivistic cultures, but the same degree of importance does not attach to group membership in individualistic cultures.

It follows that collectivistic cultures require a greater degree of harmony, cohesion, and cooperation within their in-groups and place greater burdens on individuals to identify with the group and conform to group norms. Sanctions usually exist for nonconformity. Individualistic cultures depend less on groups and more on the uniqueness of their individuals. The pursuit of personal goals rather than collective ones is of primary importance. As a result, individualistic cultures require less harmony and cohesion within groups and place less importance on conformity of individuals to group norms.

These differences in the meaning of self–in-group relationships between individualistic and collectivistic cultures have consequences for behavior. In collectivistic cultures, for example, we would expect people to make more individual sacrifices for their in-groups in pursuit of group goals. We would expect to see people trying harder to cooperate with one another, even if it means that the individual must suppress his or her own feelings, thoughts, behaviors, or goals to maintain harmony and cohesion. We would expect people to try to find ways of agreeing with each other, downplaying and minimizing interpersonal differences for the sake of harmony.

Self–in-group relationships in individualistic cultures have different consequences for behavior. In these cultures, we would expect people to make fewer sacrifices of their own individual goals, needs, and desires for the sake of a common good. We would expect people to be more expressive of their own feelings, attitudes, and opinions, without as much fear or worry about the consequences for group harmony or cohesion. We would expect people to bring up interpersonal concerns, problems, and conflicts more freely.

Not only do self–in-group relationships differ between individualistic and collectivistic cultures, but self–out-group relationships also differ. In collectivistic cultures, the primary focus of attention is on in-group relationships. For that reason, relationships with out-group people are marked by a relative lack of concern. To the degree that members of collectivistic cultures focus on harmony, cohesion, and cooperation in in-group relations, they tend to exhibit distancing, aloofness, and even discrimination with regard to self–out-group relationships. The opposite is true in individualistic cultures. People of individualistic cultures are more likely to treat out-group persons more equally, with relatively less distinction between in-groups and out-groups. Members of individualistic cultures engage in positive, relationship-building

41

behaviors with out-group others that members of collectivistic cultures would reserve only for in-group others. These concepts are summarized in Table 14.1.

Table 14.1 Self–In-group and Self–Out-group Relationship Differences as a Function of Individualism and Collectivism

Type of Culture	Characteristics
In individualistic cultures . . .	1. People have more in-groups. 2. People are not as attached to any single in-group, because there are numerous in-groups to which they can be attached. 3. Survival of the individuals and the society is more dependent on the successful and effective functioning of individuals rather than groups. 4. People make relatively fewer distinctions between in-groups and out-groups.
In collectivistic cultures . . .	1. People have fewer in-groups. 2. People are very attached to the in-groups to which they belong. 3. Survival of the individuals and the society is more dependent on the successful and effective functioning of the groups rather than individuals. 4. People make greater distinctions between in-group and out-group others.

Recent cross-cultural research on in-group and out-group relationships. Recent research continues to support many of these notions originally proposed by Triandis and his colleagues (1988). In Chapter 9, for example, we discussed two studies that found differences in emotional expressions among the United States, Japan, Poland, and Hungary that supported these ideas concerning cultural differences in self–in-group and self–out-group relationships. In another study, Goodwin and Lee (1994) asked British and Singapore Chinese university students to complete a questionnaire designed to assess the degree to which they would discuss or do 35 behaviors that are typically considered taboo. The items ranged from discussion of sexual fantasies to crying in front of friends. Participants answered these items in relation to interactions with close friends of both sexes. The results indicated that the British participants were more likely than the Chinese participants to discuss taboo subjects with their close friends. These findings are consistent with the greater maintenance of harmony and cohesion in in-group relationships in more collectivistic cultures.

In another study, Wheeler, Reis, and Bond (1989) asked participants in the United States and Hong Kong to complete the Rochester Interaction Record (RIR) for a period of two weeks. The RIR involves the self-description of every interaction that lasts for more than 10 minutes in which interactants attended to and adjusted their behavior between each other. Each interaction is described in terms of when it occurred, its length, who the interactants were, their sex breakdown, and eight scales involving ratings of disclosure, quality, satisfaction, and the nature of the interaction. The results indicated that the Hong Kong Chinese students had longer but fewer interactions, with fewer people, than did the American students. The Chinese students also reported a higher percentage of group and task interactions, and indicated great self- and other-disclosure. These findings are certainly consonant with the previous discussion concerning the type, number, and quality of self–in-group interactions in collectivistic versus individualistic cultures.

Summary

People in all cultures and societies grow up learning to make distinctions among others in terms of in-groups and out-groups. Culture exerts considerable influence not only over the structure and format of those self–in-group and self–out-group relationships but also over the very meaning of those relationships.

42

And cultural differences in the meaning of those relationships produces real, observable differences in the behaviors, thoughts, and feelings of the individual when interacting with in-group and out-group others. How we conceptualize and act on our relationships with others can differ dramatically from culture to culture. Therefore, in interacting with others, if we try to interpret their social behaviors within our own limited cultural framework, it is very possible that we will misunderstand and misinterpret those behaviors. Good intentions may be seen as bad; innocuous behaviors may be seen as threatening or aggressive.

VISION IAS

Section - B

www.numerons.com

5. Answer the following, each in not more than 150 words: **12x5=60**

(a) Discuss various dimensions of thought processes in relation to concept.

TOPIC: *Thinking and Problem Solving*
SUBTOPIC: *Concept Formation Processes*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

Concepts are cognitive elements that combine to generatively produce an infinite variety of thoughts. Just as a finite set of building blocks can be constructed into an endless variety of architectural structures, so can concepts act as building blocks for an endless variety of complex thoughts. Claiming that concepts are cognitive elements does not entail that they are primitive elements in the sense of existing without being

43

©VISION IAS
www.visionias.wordpress.com
Contact: +91-9650617807, +91-9968029039

©Numerons
www.numerons.com
Contact: +91-9560344245

learned and without being constructed from other concepts. Some theorists have argued that concepts such as **bachelor**, **kill**, and **house** are primitive in this sense (Fodor, 1975; Fodor, Garrett, Walker, & Parkes, 1980), but a considerable
body of evidence suggests that concepts typically are acquired elements that are themselves decomposable into semantic elements (McNamara & Miller, 1989).

Once a concept has been formed, it can enter into compositions with other concepts. Several researchers have studied how novel combinations of concepts are produced and comprehended. For example, how does one interpret the term buffalo paper when one first hears it? Is it paper in the shape of buffalo, paper used to wrap buffaloes presented as gifts, an essay on the subject of buffalo, coarse paper, or something like fly paper but used to catch bison? Interpretations of word combinations are often created by finding a relation that connects the two concepts. In Murphy's (1988) concept specialization model, one interprets noun-noun combinations by finding a variable that the second noun has that can be filled by the first noun. By this account, a robin snake might be interpreted as a snake that eats robins once **robin** is used to the fill the **eats** slot in the **snake** concept. Wisniewski (1997, 1998; Wisniewski & Love, 1998) has argued that properties from one concept are often transferred to another concept, and that this is more likely to occur if the concepts are similar, with parts that can be easily aligned. By this account, a robin snake may be interpreted as a snake with a red belly, once the attribute **red breast** from the robin is transferred to the snake.

In addition to promoting creative thought, the combinatorial power of concepts is required for cognitive systematicity (Fodor & Pylyshyn, 1988). The notion of systematicity is that a system's ability to entertain complex thoughts is intrinsically connected to its ability to entertain the components of those thoughts. In the field of conceptual combination, this has appeared as the issue of whether the meaning of a combination of concepts can be deduced on the basis of the meanings of its constituents. On the one hand, there are some salient violations of this type of systematicity. When adjective and noun concepts are combined, there are sometimes emergent interactions that cannot be predicted by the "main effects" of the concepts themselves. For example, the concept **gray hair** is more similar to **white hair** than to **black hair**, but **gray cloud** is more similar to **black cloud** than to **white cloud** (Medin & Shoben, 1988). Wooden spoons are judged to be fairly large (for spoons), even though this property is not generally possessed by wooden objects or spoons in general (Medin & Shoben, 1988). On the other hand, there have been notable successes in predicting how well an object fits a conjunctive description based on how well it fits the individual descriptions that comprise the conjunction (Hampton, 1987, 1997; Storms, De Boeck, Hampton, & Van Mechelen, 1999). A reasonable reconciliation of these results is that when concepts are combined the concepts' meanings systematically determine the meaning of the conjunction, but emergent interactions and real-world plausibility also shape the conjunction's meaning.

(b) Is facial expression of emotion innate or acquired? Support your answer with suitable evidence.

TOPIC: *Motivation and Emotion*
SUBTOPIC: *Psychological and Physiological Basis of Motivation and Emotion*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

A number of investigators have contributed much to the study of the facial expression of emotion: Izard, Camras and Zajonc, for example. But none has contributed more than Paul Ekman, so it is his name that heads this section. Camras, Holland and Patterson (1993) offer a very useful review of ideas in this area, and some of what follows owes a debt to their analysis.

Although Ekman and Izard do not always agree and their views on the nature of emotion differ, they both (Ekman, 1972; Izard, 1977) began by assuming that the apparently universal recognition of emotion expression depends on an innate programme for each of the primary emotions. However, more recently, they both (Ekman, 1992b; Izard, 1991) suggest that facial expressions do not always go with emotions.

Ekman (e.g., 1982, 1992b) believes that there exist three differentiated, but interrelated systems of emotion: cognition, facial expression and autonomic nervous system (ANS) activity. He admits the possibility that any aspect of emotion might be mediated by cognition, but emphasizes the significance of facial expression. Simply changing facial expression changes how one feels. Ekman stresses pattern changes in expression and physiology, arguing that language is inadequate to account for the boundaries of emotion. A particular emotion, from this perspective, might be highly differentiated in one language and entirely missing in another.

Ekman sees emotion as having 10 major characteristics:

1. There is a distinctive pan-cultural signal for each emotion;
2. There are distinctive, universal, facial expressions of emotion that can also be traced phylogenetically;
3. Emotional expression involves multiple signals;
4. The duration of emotion is limited;
5. The timing of emotional expression reflects the details of a particular emotional experience;
6. Emotional expressions can be graded in intensity, reflecting variations in the strength of the subjective experience;
7. Emotional expression can be totally inhibited;
8. Emotional expressions can be convincingly simulated;
9. Each emotion has pan-human commonalities in its elicitors;
10. Each emotion has a pan-human pattern of ans and central nervous system (cns) change.

These characteristics lead Ekman to rest his facial expression theory of emotion on three assumptions:

1. Emotion has evolved to manage the fundamental tasks of life;
2. To be adaptive, there must be a distinct pattern for each emotion;
3. Finally, there is a general coherence in that within each emotion an interconnected pattern in expression and physiology is linked to appraisal.

Ultimately, then, Ekman is emphasizing cognition.

In passing, in this consideration of Ekman's contribution, it is worth making brief mention of his work with Friesen. Ekman and Friesen (e.g., 1969) make a very influential analysis of what they termed 'non-verbal leakage' in a discussion of the importance of the body to emotional communication. They suggest that non-verbal behaviour escapes the efforts that we make at social deception (hiding our feelings) and in fact allows our real feelings to leak out. We attempt to deceive others about our feelings and we may attempt to deceive ourselves.

Ekman and Friesen characterize the deceptions as having three dimensions:

1. Saliency is the degree to which the deception is of obvious importance to the interactants, a function both of the situation and of personality.
2. The roles adopted by the interactants; for example, whether they are both deceiving and detecting, or adopting complementary roles, or whatever.
3. Collaboration of antagonism refers to an implicit pact or lack of it about the discovery and/or the continuation of the deception.

Leaving these interesting conjectures aside, the major theoretical hypotheses in this area deal with lateralization, efference and facial feedback. There are two types of lateralization hypothesis. The first is that positive emotions are mediated by the left cortical hemisphere and negative emotions by the right. And the second is that the emotions that go with approach are mediated by the left hemisphere and those that go with withdrawal by the right.

The efference view has it that the programmes for the discrete emotions produce distinct expressions through efference to the facial musculature. Camras (1991, 1992) looks at this in terms of her dynamic systems model (see Chapter 9).

The facial feedback hypothesis suggests that there might be proprioceptive, cutaneous or vascular feedback from facial expressions that influence emotional experience. The theory urges that the feedback either creates the experience or merely influences it. All these hypotheses have their support, although Camras et al. (1993) conclude that there is particularly strong support for the facial feedback hypothesis.

In their review of facial efference, Adelmann and Zajonc (1989) draw attention to a number of interesting theoretical issues. Their core question concerns how facial efference plays a causal role in the experience of emotion. Although there is little theoretical development in this area, Zajonc (1985) himself associates emotional efference to vascular systems, a notion originally put forward by Waynbaum in 1907. This is the vascular theory of emotional efference, which is based on the affect of facial muscles on venous blood flow. Zajonc sees this theory as potentially accounting for a number of apparently disparate matters such as biofeedback, placebo effects, unconscious preferences and aversions, and so on.

Adelmann and Zajonc derive various empirically testable hypotheses from the theory. However, they also conclude that there is insufficient evidence to reject any of the theories of the links between facial efference and emotional experience. In particular, they draw attention to views that point to the importance of sensory processes (e.g., Le Doux, 1987). They also point out that the facial feedback hypothesis does not cope well with why some facial expressions feel 'good' and others feel 'bad'.

In some ways, the most interesting aspect of the vascular theory of emotion that Zajonc and his co-workers espouse (see also, Zajonc, Murphy & Inglehart, 1989) is that it deals with issues that cognitive appraisal theories do not, at least as yet. In particular, the theory has it that facial expressions affect the cavernous sinus, restricting venous blood flow and thereby having an effect on the cooling of the arterial blood supply to the brain. Further, changes in the cerebral temperature could have an effect on the release or blocking of neurotransmitters that are to do with emotion. In general, this theory and some of the evidence to which it has given rise suggest possible mechanisms whereby emotion and cognition could function independently.

(c) Discuss the modifications of Chomsky's 'transformational generative grammar'.

TOPIC: *Language and Communication*
SUBTOPIC: *Theories of language development - Chomsky*
LEVEL: *Medium*
NATURE: *Fundamental*

REFERENCE:

Minimalist program

In linguistics, the Minimalist Program (MP) is a major line of inquiry that has been developing inside Generative Grammar since the early nineties. It started with a 1993 paper by Noam Chomsky.

Chomsky presents MP as a program, not as a theory, following Imre Lakatos's distinction. The MP seeks to be a mode of inquiry characterized by the flexibility of the multiple directions that its minimalism enables. Ultimately, the MP provides a conceptual framework used to guide the development of grammatical theory. For Chomsky, there are minimalist questions, but the answers can be framed in any theory. Of all these questions, the one that plays the most crucial role is this: why language has the properties it has. That said, the MP lays out a very specific view of the basis of syntactic grammar that, when compared to other formalisms, is often taken to look very much like a theory.

Theoretical goals

Perfection

The MP appeals to the idea that the language ability in humans shows signs of being incorporated under an optimal design with exquisite organization, which seems to suggest that the inner workings conform to a very simple computational law or a particular mental organ. In other words, the MP works on the assumption that Universal Grammar constitutes a perfect design in the sense that it contains only what is necessary to meet our conceptual, and physical (phonological) needs.

From a theoretical standpoint, and in the context of generative grammar, the MP draws on the minimalist approach of the Principles and Parameters program, considered to be the ultimate standard theoretical model that generative linguistics has developed since the eighties. What this approach suggests is the existence of a fixed set of principles valid for all languages, which, when combined with settings for a finite set of binary switches (parameters), may describe the specific properties that characterize the language system a child eventually comes to attain.

The MP aims to get to know how much of the Principles and Parameters model can be taken as a result of this hypothetical optimal and computationally efficient design of the human language faculty. In turn, more developed versions of the Principles and Parameters approach provide technical principles from which the MP can be seen to follow.

Economy

The MP aims at the further development of ideas involving economy of derivation and economy of representation, which had started to become significant in the early 1990s, but were still peripheral aspects of Transformational grammar.

Economy of derivation is a principle stating that movements (i.e. transformations) only occur in order to match interpretable features with uninterpretable features. An example of an interpretable feature is the plural inflection on regular English nouns, e.g. dogs. The word dogs can only be used to refer to several dogs, not a single dog, and so this inflection contributes to meaning, making it interpretable. English verbs are inflected according to the number of their subject (e.g. "Dogs bite" vs "A dog bites"), but this information is only interpretable once a relationship is formed between the subject and the verb, so movement of the subject is required.

Economy of representation is the principle that grammatical structures must exist for a purpose, i.e. the structure of a sentence should be no larger or more complex than required to satisfy constraints on grammaticality, which are equivalent to constraints on the mapping between the conceptual/intensional and sensori-motor interfaces in the optimal system that minimalism seeks to explore.

Technical innovations

The exploration of minimalist questions has led to several radical changes in the technical apparatus of transformational generative grammatical theory. Some of the most important are:

- The generalization of X-bar Theory into Bare Phrase Structure (see below).
- The simplification of representational levels in the grammatical model, eliminating the distinction between Deep Structure and Surface Structure in favor of more explicitly derivational approach.
- The elimination of the notion of government.
- The inclusion of a single point of interaction between syntax and the interfaces (conceptual/intensional and sensori-motor), commonly called the point of Spell-Out.
- The idea that syntactic derivations proceed by clearly delineated stages called phases (see below).

Bare Phrase Structure

A major development of MP inquiry is Bare Phrase Structure (BPS), a theory of phrase structure (sentence building prior to movement) developed by Noam Chomsky.

This theory contrasts with X-bar theory, which preceded it, in four important ways:

- BPS is explicitly derivational. That is, it is built from the bottom up, bit by bit. In contrast, X-Bar Theory is representational - a structure for a given construction is built in one fell swoop, and lexical items are inserted into the structure.
- BPS does not have a preconceived phrasal structure, while in X-Bar Theory, every phrase has a specifier, a head, and a complement.
- BPS permits only binary branching, while X-Bar Theory permits both binary and unary branching.
- BPS does not distinguish between a "head" and a "terminal", while some versions of X-Bar Theory require such a distinction.

BPS incorporates two basic operations: Merge and Move. Although there is active debate on exactly how Move should be formulated, the differences between the current proposals are relatively minute. The following description follows Chomsky's original proposal.

Merge is a function that takes two objects (say α and β) and merges them into an unordered set with a label (either α or β, in this case α). The label identifies the properties of the phrase.

Merge (α, β) → {α, {α, β}}

For example, Merge can operate on the lexical items 'drink' and 'water' to give 'drink water'. Note that the phrase 'drink water' behaves more like the verb 'drink' than like the noun 'water'. That is, wherever we can put the verb 'drink' we can usually put the phrase 'drink water':

I like to _____ (drink)/(drink water).
(Drinking/Drinking water) _____ is fun.

Furthermore, we typically can't put the phrase 'drink water' in places where we can put the noun 'water':

We can say "There's some water on the table", but not "There's some drink water on the table".

So, we identify the phrase with a label. In the case of 'drink water', the label is 'drink' since the phrase acts as a verb. For simplicity, we call this phrase a verb phrase or VP. Now if we were to Merge 'cold' and 'water' to get 'cold water', then we would have a noun phrase or NP with the label 'water'. The reader can verify that the phrase 'cold water' can appear in the same environments as the noun 'water' in the three test sentences above. So, for 'drink water' we have the following:

Merge (drink, water) → {drink, {drink, water}}

We can represent this in a typical syntax tree as follows:

```
    drink
    /  \
drink  water
```

or, with more technical terms, as:

```
    VP
    / \
drink  water
```

Merge can also operate on structures already built. If it couldn't, then such a system would predict only two-word utterances to be grammatical. Say we Merge a new head with a previously formed object (a phrase).

Merge (γ, {α, {α, β}}) → {γ, {γ, {α, {α, β}}}}

Here, γ is the label, so we say that γ 'projects' from the label of the head. This corresponds to the following tree structure:

```
 γ
 / \
γ  α
  / \
 α  β
```

Note crucially that Merge operates blindly, projecting labels in all possible combinations. The subcategorization features of the head then license certain label projections and eliminate all derivations with alternate projections.

Phases

A phase is a syntactic domain first hypothesized by Noam Chomsky in 1998. A simple sentence is often decomposed into two phases, CP and vP (see X-bar theory). Movement of a constituent out of a phase is (in the general case) only permitted if the constituent has first moved to the left edge of the phase. This condition is described in the Phase Impenetrability Condition, which has been variously formulated within the literature. In its original conception, only the vP in transitive and unergative verbs constitute phases. The vP in passives and unaccusative verbs (if even present) are not phases. This topic is, however, currently under debate in the literature.

Criticism

In the late 1990s, David E. Johnson and Shalom Lappin published the first detailed critiques of Chomsky's minimalist program. This technical work was followed by a lively debate with proponents of minimalism on the scientific status of the program. The original article provoked several replies and two further rounds of replies and counter-replies in subsequent issues of the same journal. Lappin et al. argue that the Minimalist Program is a radical departure from earlier Chomskian linguistic practice that is not motivated by any new empirical discoveries, but rather by a general appeal to "perfection" which is both empirically unmotivated and so vague as to be unfalsifiable. They compare the adoption of this paradigm by linguistic researchers to other historical paradigm shifts in natural sciences and conclude that the adoption of the Minimalist Program has been an "unscientific revolution", driven primarily by Chomsky's authority in linguistics. The several replies to the article in Natural Language and Linguistic Theory Volume 18 number 4 (2000) make a number of different defenses of the Minimalist Program. Some claim that it is not in fact revolutionary or not in fact widely adopted, while others agree with Levine and Johnson on these points, but defend the vagueness of its formulation as not problematic in light of its status as a research program rather than a theory (see above).

(d) Compare LISREL programme with that of SPSS in the analysis of multivariate data.

TOPIC: Research Methods
SUBTOPIC: Application of statistical technique
LEVEL: Difficult
NATURE: Unconventional

REFERENCE:

Introduction

The software package LISREL provides calculations for Structural Equations Modeling, mostly multivariate analysis based on regressions. This includes Factor Analysis, Path Analysis, Multilevel Modelling, and other

complex statistical procedures related to multivariate measurements. This web page is concerns only with Factor Analysis.

There are two major types of Factor Analysis, exploratory and confirmatory. Exploratory Factor Analysis are explained in Factor_Exp.php. The remainder of this page shall describe Confirmatory Factor Analysis, and presents code samples on how to do both in LISREL.

LISREL

LISREL is a software package produced by Scientific Software International (SSI) . It started off as a package for statisticians, where the control of the programs and the format of graphical output was quite challenging, particularly to non-statisticians. More recently, SSI develop its new control language Simplis, which is more like ordinary English. With this, LISREL became very easy to use.

When using LISREL, the control of the program is by means of a text file with the .spl extension. When this is run, the program output is in a file with the .out extension. If there is a path diagram, this is in a file with .pth extension.

LISREL boasts its path diagram, saying this is publication quality. This is only true when there are few variables, such as in student exercises. In most research work, may variables are used, and the path diagram becomes too large and unweildy when more than 10 variables are used. My advice is to rely on the .out file only and draw your own path diagrams.

This page presents templates for both exploratory and confirmatory factor analysis using LISREL, so that users can easily adapt them for their own work.

Exploratory Factor Analysis

Exploratory Factor Analysis is usually carried out for two purposes. The first is to explore the relationships between a set of measurements or variables, and the second is a mean of data reduction, where a large number of variables can be summarised or compressed into a fewer number of vectors.

Exploratory factor analysis has no apriori theory or hypothesis, and is sometimes call unsupervised clustering. The variables are clustered according to how they correlated with each other.

LISREL carries out exploratory factor analysis in a slightly different manner to SPSS.

In SPSS, the data input is usually the raw data, the default extraction method is the least squares, and the number of factor is determined by the scree test, stopping when the communality falls below 1. The type of rotation is then stipulated.

In LISREL, input can be the raw data, but it will also accepts the covariance or the correlation matrix.

The default extraction method is Maximum Likelihood, where factors are extracted until the residual matrix contains no remaining correlation (Chi square of the residual matrix no longer statistically significant). This is a much more thorough method, and usually will extract more factors than that using other methods.

The problem with maximum likelihood is that the Heywood situation may occur. This is when all of the variance in one or more variables are exhausted while there are still correlations left in the overall matrix. This occurs particularly if the strength of the correlation coefficient varies widely in the correlation matrix. When the Heywood situation occurs, the results are distorted and can no longer be considered valid. The factor analysis will then have to be repeated specifying a reduced number of factors until the Heywood situation is avoided.

If there are great discrepancies in the strength of the correlation coefficients in the data, factor extraction may be caught between the Heywood situation if too many factors are extractede, and a failure to exhaust

all correlations in the matrix if insufficient factors are extracted. If this occurs, the factor matrix will not account for all the correlations available in the input matrix, and the Chi Square Test on the residual matrix will remain statistically significant.

If statistical robustness is not a major priority, such as is the case for most data reduction exercises, extraction using MINRES can be specified. This will produce a factor matrix very similar to that from SPSS (using default setting).

The output of LISREL can be specified. If unsure, all can be specified, and everything comes out, pages of matrices with very little explanations. My advice is to look at only the matrices after Varimax and Promax rotation, and the correlation between the factors.

Confirmatory Factor Analysis

Confirmatory factor analysis answers the question whether a set of data fits a prescribed factor pattern. It is usually used for two purposes.

The first is to test or confirm that a factor pattern, say from a survey tool, is stable and therefore be confirmed by an independent set of data. In this, the factor pattern comes from an existing tool or a theoretical construct, and whether this fits with a set of data is then tested.

The second is in the development of a multivariate instrument or tool, such as questionnaire to evaluate racism. In this, the number of factors (concepts, constructs, or dimensions) are firstly defined, then a number of variables (questions or measurements) that may reflect each of these construct developed. Data of are then collected, and tested against the factor-variable relationship. Those variables that do not fit neatly into a single factor are then replaced or changed, and new data are collected and tested. This process is repeated until a set of data collected fits the required pattern.

Confirmatory factor analysis uses the Maximum Likelihood method of extraction, so it has a degree of statistical robustness, and allows for significance testing. In practice, however, statistical significance is difficult to interpret, as it is determined not only by how well the data fits in with the theoretical construct, but also by sample size, and the number of variables and factors.

Test of fit between data and construct

The Chi Square Test is the primary statistical test. However, Chi Squares tends to increase with sample size, with decreasing number of variables, and increasing number of factors. So, unless the theoretical construct was developed using Maximum Likelihood Factor Analysis that had exhausted all the correlations in the original matrix, and that the sample size of the testing data set is similar to the original data set used to develop the construct, the Chi Square may not truly reflect how well the data fits the construct.

As the Chi Square Test is statistical robust, yet problematical because of the complexity of confirmatory factor analysis, statisticians have developed an array of adaptation of the Chi Square to adjust for sample size and the relationship between the number of variables and factors. Appendix 3 present a synopsis of the indices produced by LISREL, and whether they are affected by sample size.

From the clinical point of view however, the following decision making steps are recommended by some of the publications.

- Step 1. Examine the critical number. This is the minimum sample size, below which the results cannot be validly interpreted. Only if the sample size exceeds the critical number can interpretation proceed.
- Step 2. Look at the significance of the Chi Square. The Minimum Fit Chi Square can be used if one can be sure that all the variables are continuous measurement and normally distributed. If this cannot be assured, as is the case in almost all cases, then the Normal Theory Weighted Least Squares Chi-Square is used. If the Chi Square is not significant ($p >= 0.05$), then a decision that a

51

©VISION IAS
www.visionias.wordpress.com
Contact: +91-9650617807, +91-9968029039

©Numerons
www.numerons.com
Contact: +91-9560344245

good fit exists between data and theory can be made. Subsequent steps are only necessary of the Chi Squares is significant ($p < 0.05$) .

- Step 3. Examine the Chi Square Degree of Freedom Ratio. This is obtained Ratio = Chi Sq / Deg Freedom. If the Ratio is 2 or less, then a decision that a reasonably good fit exists can be made. If the ratio exceeds 2 go to step 4.
- Step 4. Examine the Goodness of Fit Index. If this index is 0.9 or more, then a decision that a reasonable good fit exists can be made. If not then the conclusion that the data fits poorly to theory should now be made.
- Possible actions when data and theory do not fit.

If the conclusion is that the data and theory fits, then the statistical exercise ends. The researcher accepts the validity of the theory and moves on.

When the conclusion is that the data do not fit the theory, the actions that are possible then depends on the reasons for conduction the confirmatory test in the first place.

Option 1. Reject the factor construct and move on. This is of course the primary purpose of the exercise. The question is whether the data fits the theory, the answer is no. It ends. Another reason for doing this is if the data itself is problematical, so that a fit can never be obtained. This occurs if the number of variables in a factor are too few, if the correlations are such that one or more variables load across more than 1 factor, or if the correlations within a factor vary greatly (as it would then not be possible to extract all the correlation out of the matrix).

Option 2. Why did it not fit and is it fixable. This option can be taken if the fit is nearly good enough, and only a few variables are problematic. The table Modification Indices for LAMBDA-X (example in appendix 3.c.) can be examined. This is the table of Chi Squares that can be reduced if an additional link between a factor and a variable is made. All Chi Square values exceeding 5 are considered excessive.

The approach is to add the link with the largest Chi Square value to the spl file under the command keyword Relationship, and re-run the confirmatory factor analysis. If there is still no fit, then add the link with the next biggest chi square, and re-run. Keep repeating this until a fit is found.

There are two problems that may arise. The first is that, if more than 2-3 additional links are put in, the factor model becomes confusing and theoretically untenable. The second is that, with the addition of a link, all other relationships adjusted a bit, and the relationships may be such that any improvement on one part of the model causes a deterioration on another part, and things just do not get any better.

Option 3. Change the variables. This is only possible if confirmatory factor analysis is used during the development of a new multivariate instrument. In this situation, a model will not fit if the number of variables that link to a factor are too few (5-7 are recommended), or a variable links significantly to more than 1 factor. Confirmatory factor analysis will identify these, and the thing to do is to increase, modify or replace the appropriate variables and try again, until all variables only link to single factors. When this happens an acceptable fit will result.

(e) Evaluate various factors which influence intrinsic motivation.

TOPIC: *Motivation and Emotion*
SUBTOPIC: *Extrinsic and Intrinsic Motivation*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

INTRINSIC MOTIVATION

Intrinsic motivation is defined as the doing of an activity for its inherent satisfactions rather than for some separable consequence. When intrinsically motivated a person is moved to act for the fun or challenge entailed rather than because of external prods, pressures, or rewards. The phenomenon of intrinsic motivation was first acknowledged within experimental studies of animal behavior, where it was discovered that many organisms engage in exploratory, playful, and curiosity-driven behaviors even in the absence of reinforcement or reward (White, 1959). These spontaneous behaviors, al- though clearly bestowing adaptive benefits on the organism, appear not to be done for any such instrumental reason, but rather for the positive experiences associated with exercising and extending ones capacities.

In humans, intrinsic motivation is not the only form of motivation, or even of volitional activity, but it is a pervasive and important one. From birth onward, humans, in their healthiest states, are active, inquisitive, curious, and playful creatures, displaying a ubiquitous readiness to learn and explore, and they do not require extraneous incentives to do so. This natural motivational tendency is a critical element in cognitive, social, and physical development because it is through acting on one's inherent interests that one grows in knowledge and skills. The inclinations to take interest in novelty, to actively assimilate, and to creatively apply our skills is not limited to childhood, but is a significant feature of human nature that affects performance, persistence, and well-being across life's epochs (Ryan & LaGuardia, in press).

Although, in one sense, intrinsic motivation exists within individuals, in another sense intrinsic motivation exists in the relation between individuals and activities. People are intrinsically motivated for some activities and not others, and not everyone is intrinsically motivated for any particular task.

Because intrinsic motivation exists in the nexus between a person and a task, some authors have defined intrinsic motivation in terms of the task being interesting while others have defined it in terms of the satisfactions a person gains from intrinsically motivated task engagement. In part, these different definitions derive from the fact that the concept of intrinsic motivation was proposed as a critical reaction to the two behavioral theories that were dominant in empirical psychology from the 1940s to the 1960s.

Specifically, because operant theory (Skinner, 1953) maintained that all behaviors are motivated by rewards (i.e., by separable consequence such as food or money), intrinsically motivated activities were said to be ones for which the reward was in the activity itself. Thus, researchers investigated what task characteristics make an activity interesting. In contrast, because learning theory (Hull, 1943) asserted that all behaviors are motivated by physiological drives (and their derivatives), intrinsically motivated activities were said to be ones that provided satisfaction of innate psychological needs. Thus, researchers explored what basic needs are satisfied by intrinsically motivated behaviors.

Our own approach focuses primarily on psychological needs—namely, the innate needs for competence, autonomy, and relatedness—but we of course recognize that basic need satisfaction accrues in part from engaging in interesting activities. Thus, we do sometimes speak of intrinsically inter- esting activities, but when we do so we are really only talking about tasks that, on average, many people find to be intrinsically interesting. There is considerable practical utility in focusing on task properties and their potential intrinsic interest, as it leads toward improved task design or selection to enhance motivation.

Operational Definitions

Intrinsic motivation has been operationally defined in various ways, al- though there have been two measures that have been most often used. Basic experimental research (e.g., Deci, 1971) has rested primarily on a behavioral measure of intrinsic motivation called the ''free choice'' measure. In experiments using this measure participants are exposed to a task under varying conditions (e.g., getting a reward or not). Following this period, the experimenter tells participants they will not be asked to work with the target task any further, and they are then left alone in the experimental room with the target task as well as various distractor activities. They thus have a period of ''free choice'' about whether to return to the activity, and it is assumed that, if there is no extrinsic reason to do the task (e.g., no reward and no approval), then the more time they spend with the target task, the more intrinsically motivated they are for that task. This measure has been the mainstay through which the dynamics of intrinsic motivation have been experimentally studied.

The other common approach to the measurement of intrinsic motivation is the use of self-reports of interest and enjoyment of the activity per se. Experimental studies typically rely on task-specific measures (e.g. Ryan, 1982; Harackiewicz, 1979). Most field studies have instead used more general, "domain" focused measures, such as one's intrinsic motivation for school (e.g., Harter, 1981).

Facilitating versus Undermining Intrinsic Motivation

Despite the observable evidence that humans are liberally endowed with intrinsic motivational tendencies, this propensity appears to be expressed only under specifiable conditions. Research into intrinsic motivation has thus placed much emphasis on those conditions that elicit, sustain, and enhance this special type of motivation versus those that subdue or diminish it. Self-Determination Theory is specifically framed in terms of social and environ- mental factors that facilitate versus undermine intrinsic motivation. This language reflects the assumption that intrinsic motivation, being an inherent organismic propensity, is catalyzed (rather than caused) when individuals are in conditions that conduce toward its expression.

Cognitive Evaluation Theory (CET) was presented by Deci and Ryan (1985) to specify the factors in social contexts that produce variability in intrinsic motivation. CET, which is considered a sub-theory of self-determination theory, argues that interpersonal events and structures (e.g., rewards, communications, feedback) that conduce toward feelings of competence during action can enhance intrinsic motivation for that action because they allow satisfaction of the basic psychological need for competence. Accordingly, for example, optimal challenges, effectance promoting feedback, and freedom from demeaning evaluations are all predicted to facilitate intrinsic motivation.

CET further specifies that feelings of competence will not enhance intrinsic motivation unless they are accompanied by a sense of autonomy or, in attributional terms, by an internal perceived locus of causality (IPLOC; de-Charms, 1968). Thus, people must not only experience perceived competence (or self-efficacy), they must also experience their behavior to be self-determined if intrinsic motivation is to be maintained or enhanced. Stated differently, for a high level of intrinsic motivation people must experience satisfaction of the needs both for competence and autonomy. Much of the research has focused on the effects of immediate contextual conditions that either support or thwart the needs for competence and autonomy, but some has recognized that the supports can, to some extent, come from individuals' abiding inner resources that support their ongoing feelings of competence and autonomy.

The tenets of CET, with their primary focus on the needs for competence and autonomy, were formulated to integrate a set of results from initial stud- ies of the effects of rewards, feedback, and other external events on intrinsic motivation. Subsequently, they have been confirmed in both laboratory experiments and applied field studies, many of which have been done in classrooms.

Several early studies showed that positive performance feedback enhanced intrinsic motivation (e.g., Deci, 1971; Harackiewicz, 1979), whereas negative performance feedback diminished it (e.g., Deci & Cascio, 1972). Others (e.g., Vallerand & Reid, 1984) showed that perceived competence mediated these effects, and still others supported the hypothesis that increases in perceived competence must be accompanied by a sense of autonomy in order for the enhanced feelings of competence to result in increased intrinsic motivation (Ryan, 1982).

In fact, the majority of the research on the effects of environmental events on intrinsic motivation has focused on the issue of autonomy versus control rather than that of competence. And this issue has been considerably more controversial. The research began with the demonstration that extrinsic re- wards can undermine intrinsic motivation (Deci, 1971; Lepper, Greene, & Nisbett, 1973), which we interpret in terms of the reward shifting people from a more internal to external perceived locus of causality. Although the issue of rewards has been hotly debated, a recent meta-analysis (Deci, Koestner, & Ryan, in press) confirms that virtually every type of expected tangible reward made contingent on task performance does, in fact, undermine intrinsic motivation. Furthermore, not only tangible rewards, but also threats (Deci & Cascio, 1972), deadlines (Amabile, DeJong, & Lepper, 1976), directives (Koestner, Ryan, Bernieri, & Holt, 1984), and competition pressure (Reeve & Deci, 1996) diminish intrinsic motivation because, according to CET,

people experience them as controllers of their behavior. On the other hand, choice and the opportunity for self-direction (e.g., Zuckerman, Porac, Lathin, Smith, & Deci, 1978) appear to enhance intrinsic motivation, as they afford a greater sense of autonomy.

The significance of autonomy versus control for the maintenance of intrinsic motivation has been clearly observed in studies of classroom learning. For example, several studies have shown that autonomy-supportive (in con- trast to controlling) teachers catalyze in their students greater intrinsic motivation, curiosity, and the desire for challenge (e.g., Deci, Nezlek, & Sheinman, 1981; Ryan & Grolnick, 1986). Students who are overly controlled not only lose initiative but also learn less well, especially when learning is complex or requires conceptual, creative processing (Benware & Deci, 1984; Grolnick & Ryan, 1987). Similarly, studies show children of parents who are more autonomy supportive to be more mastery oriented—more likely to spontaneously explore and extend themselves—than children of parents who are more controlling (Grolnick, Deci, & Ryan, 1997).

To summarize, the CET aspect of SDT suggests that classroom and home environments can facilitate or forestall intrinsic motivation by supporting versus thwarting the needs for autonomy and competence. However, it is critical to remember that intrinsic motivation will occur only for activities that hold intrinsic interest for an individual—those that have the appeal of novelty, challenge, or aesthetic value for that individual. For activities that do not hold such appeal, the principles of CET do not apply. To understand the motivation for activities that are not experienced as inherently interesting, we need to look more deeply into the nature and dynamics of extrinsic motivation.

6. Answer the following, each in not more than 250 words: 20x3=60

(a) Compare 'multi-factor' theory of intelligence with that of theory of 'multiple intelligence'.

TOPIC: *Intelligence and Aptitude*
SUBTOPIC: *Nature and Theories of Intelligence*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

MULTI-FACTOR THEORY OF INTELLIGENCE

Cattell-Horn-Carroll (Chc) Theory of Intelligence

Cattell-Horn-Carroll (CHC) theory: A multifactor theory of intelligence that integrates the Cattell-Horn Gf-Gc theory and the Carroll Three-Sratum theory. It is composed of 10 broad and more than 70 narrow cognitive abilities/processes (Flanagan et al., 2008; McGrew, 2005). Recent advances in current theory and research on the structure of human cognitive abilities have resulted in a new empirically derived model commonly referred to as the Cattell–Horn–Carroll theory of cognitive abilities (CHC theory). CHC theory of cognitive abilities is an amalgamation of two similar theories about the content and structure of human cognitive abilities. The first of these two theories is Gf-Gc theory (Raymond Cattell, 1941; Horn 1965), and the second is John Bissell Carroll's (1993) Three-Stratum theory. Carroll's expansion of Gf-Gc theory to CHC theory was developed in the course of a major survey of research over the past 60 or 70 years on the nature, identification, and structure of human cognitive abilities. That research involved the use of the mathematical technique known as factor analysis. In comparison to other well-known theories of intelligence and cognitive abilities, CHC theory is the most comprehensive and empirically supported psychometric theory of the structure of cognitive and academic abilities.

The CHC model was expanded by McGrew (1997), later revised with the help of Flanagan (1998), and extended again by McGrew (2011). There are a fairly large number of distinct individual differences in cognitive ability, and CHC theory holds that the relationships among them can be derived by classifying them into three different strata: stratum I, "narrow" abilities; stratum II, "broad abilities"; and stratum III, consisting of a single "general" ability (or g).

Briefly, research has consistently shown that intellectual functions can be clustered in several different domains. In terms of learning, each of these abilities has a particular impact on the individual's ability to learn mathematics, read, write, etc. It then follows that we need to understand how the child/individual performs in these areas in order to understand how he or she learns. Nine of the factors of intelligence listed in CHC theory are:

1. **Comprehension-knowledge (Gc)** is defined as the depth and breadth of knowledge and skills that are valued by one's culture. Comprehension/ knowledge reflects the degree to which a person has learned practically useful knowledge and mastered valued skills. It includes, among others, knowledge funds, language development, listening ability, and communication ability.

2. **Fluid intelligence/reasoning (Gf)** is the deliberate but flexible control of attention to solve novel "on the spot" problems that cannot by performed by relying exclusively on previously learned information. Its purpose is to solve unfamiliar problems. It includes, among others, inferential reasoning, concept formation, classification of unfamiliar stimuli, generalization of old solutions to new problems, hypothesis generation and testing, and extrapolation of reasonable estimates in equivocal situations. Quantitative, induction (observe and discover underlying principles), and sequential (ability to reason logically, to apply rules) reasoning are part of this construct.

3. **Short-term memory (Gsm)** is the ability to encode, maintain, and manipulate information in one's immediate awareness. It refers to both the size of primary memory and to the efficiency of attentional control mechanisms that manipulate information within primary memory. Working memory is the ability to direct the focus of attention to perform relatively simple manipulations, combinations, and transformations of information within primary memory while avoiding distracting stimuli and engaging in strategic/controlled searches for information in secondary memory.

4. **Long-term storage and retrieval (Glr)** refers to the ability to store information and fluently retrieve it later through association. It is the ability to memorize information and to retrieve it. Rapid language retrieval is part of this construct.

5. **Processing speed (Gs)** refers to the ability quickly perform automatic, routine cognitive tasks, particularly when pressed to maintain focused concentration. It is the speed with which we process rote information.

6. **Visual processing (Gv)** refers to the ability to analyze and synthesize visual stimuli. It is the ability to perceive and remember visual input.

7. **Auditory processing (Ga)** refers to the ability to analyze and synthesize auditory stimuli. It involves perception of the individual sounds that form a language.

8. **Reading & Writing Ability (Grw)**: includes basic reading and writing skills.

9. **Quantitative Reasoning (Gq)**: is the ability to comprehend quantitative concepts and relationships and to manipulate numerical symbols.

THEORY OF MULTIPLE INTELLIGENCES by HOWARD GARDNER

The theory of multiple intelligences was proposed by Howard Gardner in 1983 as a model of intelligence that differentiates intelligence into various specific (primarily sensory) modalities, rather than seeing it as dominated by a single general ability.

Gardner's theory of intelligence is based on a neuropsychological analysis of human abilities (Gardner, 1983). It argues that intelligence falls into seven categories: linguistic intelligence, musical intelligence,

logical/mathematical intelligence, spatial intelligence, bodily/kinesthetic intelligence, and two types of personal intelligence. Bodily/kinesthetic intelligence includes the types of skill that athletes, typists, dancers or mime artists exhibit. Personal intelligence includes awareness of one's own feelings (intrapersonal intelligence) and the ability to notice individual differences in other people and to respond appropriately to them – in other words, to be socially aware (interpersonal intelligence).

Three of Gardner's types of intelligence – verbal intelligence, logical/mathematical intelligence and spatial intelligence – are not unusual, having been identified previously by many other researchers. The other four are rather unusual. According to Gardner, all seven abilities are well represented in the brain, in that specific brain damage can impair some of them but leave others relatively intact. For example, people with damage to the left parietal lobe can show apraxia, an inability to perform sequences of voluntary skilled movements. In contrast, people with damage to the right parietal lobe develop the spatial neglect described in Chapter 6. Individuals with frontal lobe damage, as you saw in Chapter 4 and will see later in this chapter and in Chapter 13, have difficulty evaluating the significance of social situations and making decisions about social matters (the frontal lobes used to be regarded as the region of the brain responsible for intelligence). These examples illustrate bodily/kinesthetic intelligence and both intrapersonal and interpersonal intelligence.

Gardner's theory has the advantage of being based on neuropsychological reality. It also accommodates the views of intelligence held by some non-Western cultures. For example, he would recognize the ability of a member of the Puluwat culture of the Caroline Islands to navigate across the sea by the stars as an example of intelligence.

Gardner argues that there is a wide range of cognitive abilities, and that there are only very weak correlations among them. For example, the theory predicts that a child who learns to multiply easily is not necessarily generally more intelligent than a child who has more difficulty on this task. The child who takes more time to master simple multiplication 1) may best learn to multiply through a different approach, 2) may excel in a field outside of mathematics, or 3) may even be looking at and understanding the multiplication process at a fundamentally deeper level, or perhaps as an entirely different process. Such a fundamental understanding can result in what looks like slowness and can hide a mathematical intelligence potentially higher than that of a child who quickly memorizes the multiplication table despite possessing a less detailed understanding of the process of multiplication.

The theory has been met with mixed responses. Traditional intelligence tests and psychometrics have generally found high correlations between different tasks and aspects of intelligence, rather than the low correlations which Gardner's theory predicts. Nevertheless many educationalists support the practical value of the approaches suggested by the theory.

The Multiple Intelligences

Gardner articulated several criteria for a behavior to be an intelligence. These were that the intelligences:

- Potential for brain isolation by brain damage,
- Place in evolutionary history,
- Presence of core operations,
- Susceptibility to encoding (symbolic expression),
- A distinct developmental progression,
- The existence of savants, prodigies and other exceptional people,
- Support from experimental psychology and psychometric findings.

Gardner believes that eight abilities meet these criteria:

1. Spatial
2. Linguistic
3. Logical-mathematical
4. Bodily-kinesthetic

57

©VISION IAS
www.visionias.wordpress.com
Contact: +91-9650617807, +91-9968029039

©Numerons
www.numerons.com
Contact: +91-9560344245

5. Musical
6. Interpersonal
7. Intrapersonal
8. Naturalistic

He considers that existential and moral intelligence may also be worthy of inclusion.

The first three are closely linked to fluid ability, and the verbal and spatial abilities that form the hierarchical model of intelligence

1. **Logical-mathematical:** This area has to do with logic, abstractions, reasoning and numbers and critical thinking. While it is often assumed that those with this intelligence naturally excel in mathematics, chess, computer programming and other logical or numerical activities, a more accurate definition places less emphasis on traditional mathematical ability and more on reasoning capabilities, recognizing abstract patterns, scientific thinking and investigation and the ability to perform complex calculations. Logical reasoning is closely linked to fluid intelligence and to general ability (g-factor).

2. **Spatial:** This area deals with spatial judgement and the ability to visualize with the mind's eye. Careers which suit those with this type of intelligence include artists, designers and architects. A spatial person is also good with puzzles. Spatial ability is one of the three factors beneath g in the hierarchical model of intelligence.

3. **Linguistic:** This area has to do with words, spoken or written. People with high verbal-linguistic intelligence display a facility with words and languages. They are typically good at reading, writing, telling stories and memorizing words along with dates. They tend to learn best by reading, taking notes, listening to lectures, and by discussing and debating about what they have learned. Those with verbal-linguistic intelligence learn foreign languages very easily as they have high verbal memory and recall, and an ability to understand and manipulate syntax and structure. Verbal ability is one of the most g-loaded abilities.

4. **Bodily-kinesthetic:** The core elements of the bodily-kinesthetic intelligence are control of one's bodily motions and the capacity to handle objects skillfully. Gardner elaborates to say that this intelligence also includes a sense of timing, a clear sense of the goal of a physical action, along with the ability to train responses so they become like reflexes. In theory, people who have bodily-kinesthetic intelligence should learn better by involving muscular movement (e.g. getting up and moving around into the learning experience), and are generally good at physical activities such as sports or dance. They may enjoy acting or performing, and in general they are good at building and making things. They often learn best by doing something physically, rather than by reading or hearing about it. Those with strong bodily-kinesthetic intelligence seem to use what might be termed "muscle memory", drawing on it to supplement or in extreme cases even substitute for other skills such as verbal memory. Careers that suit those with this intelligence include: athletes, pilots, dancers, musicians, actors, surgeons, builders, police officers, and soldiers. Although these careers can be duplicated through virtual simulation, they will not produce the actual physical learning that is needed in this intelligence.

5. **Musical:** This area has to do with sensitivity to sounds, rhythms, tones, and music. People with a high musical intelligence normally have good pitch and may even have absolute pitch, and are able to sing, play musical instruments, and compose music. Since there is a strong auditory component to this intelligence, those who are strongest in it may learn best via lecture. Language skills are typically highly developed in those whose base intelligence is musical. In addition, they will sometimes use songs or rhythms to learn. They have sensitivity to rhythm, pitch, meter, tone, melody or timbre. Careers that suit those with this intelligence include instrumentalists, singers, conductors, disc jockeys, orators, writers and composers. Research measuring the effects of music on second language acquisition is supportive of this music-language connection. In an investigation conducted on a group of elementary-aged English language learners, music facilitated their

58

language learning. Gardner's theory may help to explain why music and its sub-components (i.e., stress, pitch, rhythm) may be viable vehicles for second language learning.

6. **Interpersonal:** This area has to do with interaction with others. Interpersonal intelligence is the ability to understand others. In theory, individuals who have high interpersonal intelligence are characterized by their sensitivity to others' moods, feelings, temperaments and motivations, and their ability to cooperate in order to work as part of a group. According to Gardner in How Are Kids Smart: Multiple Intelligences in the Classroom, "Inter- and Intra- personal intelligence is often misunderstood with being extroverted or liking other people..." Interpersonal intelligence means that you understand what people need to work well. Individuals with this intelligence communicate effectively and empathize easily with others, and may be either leaders or followers. They typically learn best by working with others and often enjoy discussion and debate. Careers that suit those with this intelligence include sales, politicians, managers, teachers, counselors and social workers.

7. **Intrapersonal:** This area has to do with introspective and self-reflective capacities. This refers to having a deep understanding of the self; what your strengths/ weaknesses are, what makes you unique, being able to predict your own reactions/emotions. Philosophical and critical thinking is common with this intelligence. Many people with this intelligence are authors, psychologists, counselors, philosophers, and members of the clergy.

8. **Naturalistic:** This area has to do with nurturing and relating information to one's natural surroundings. Examples include classifying natural forms such as animal and plant species and rocks and mountain types; and the applied knowledge of nature in farming, mining, etc. Careers which suit those with this intelligence include naturalists, farmers and gardeners.

9. **Existential:** Some proponents of multiple intelligence theory proposed spiritual or religious intelligence as a possible additional type. Gardner did not want to commit to a spiritual intelligence, but suggested that an "existential" intelligence may be a useful construct. The hypothesis of an existential intelligence has been further explored by educational researchers. Ability to contemplate phenomena or questions beyond sensory data, such as the infinite and infinitesimal. Careers or callings which suit those with this intelligence include shamans, priests, mathematicians, physicists, scientists, cosmologists, psychologists and philosophers.

(b) Enumerate different stages of sleep. Discuss the effects of sleep deprivation on cognitive performance.

TOPIC: *Issues and Perspectives in Modern Contemporary Psychology*
SUBTOPIC: *Study of Consciousness-Sleep-Wake Schedules*
LEVEL: *Easy*
**NATURE:**   *Fundamental*

REFERENCE:

Sleep

Sleep is a naturally recurring state characterized by reduced or absent consciousness, relatively suspended sensory activity, and inactivity of nearly all voluntary muscles. It is distinguished from quiet wakefulness by a decreased ability to react to stimuli, and is more easily reversible than being in hibernation or a coma. Sleep is also a heightened anabolic state, accentuating the growth and rejuvenation of the immune, nervous, skeletal and muscular systems. It is observed in all mammals, all birds, and many reptiles, amphibians, and fish.

The purposes and mechanisms of sleep are only partially clear and are the subject of intense research. Sleep is often thought to help conserve energy, but actually decreases metabolism only about 5–10%. Hibernating animals need to sleep despite the hypometabolism seen in hibernation, and in fact they must return from hypothermia to euthermy in order to sleep, making sleeping "energetically expensive."

Physiology

Sleep stages

In mammals and birds, sleep is divided into two broad types: rapid eye movement (REM) and non-rapid eye movement (NREM or non-REM) sleep. Each type has a distinct set of associated physiological, neurological, and psychological features. The American Academy of Sleep Medicine (AASM) further divides NREM into three stages: N1, N2, and N3, the last of which is also called delta sleep or slow-wave sleep (SWS).
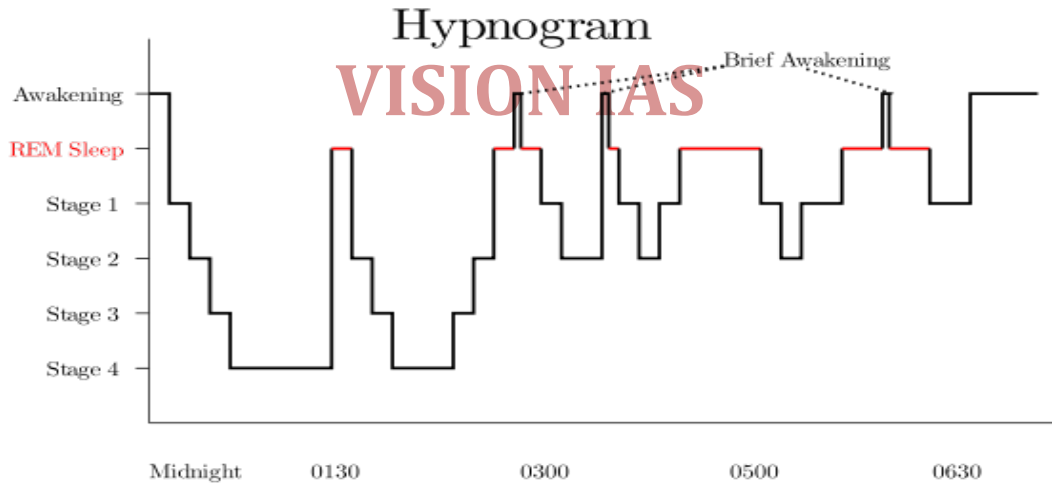
Hypnogram showing sleep cycles from midnight to 6.30 am, with deep sleep early on. There is more REM (marked red) before waking.
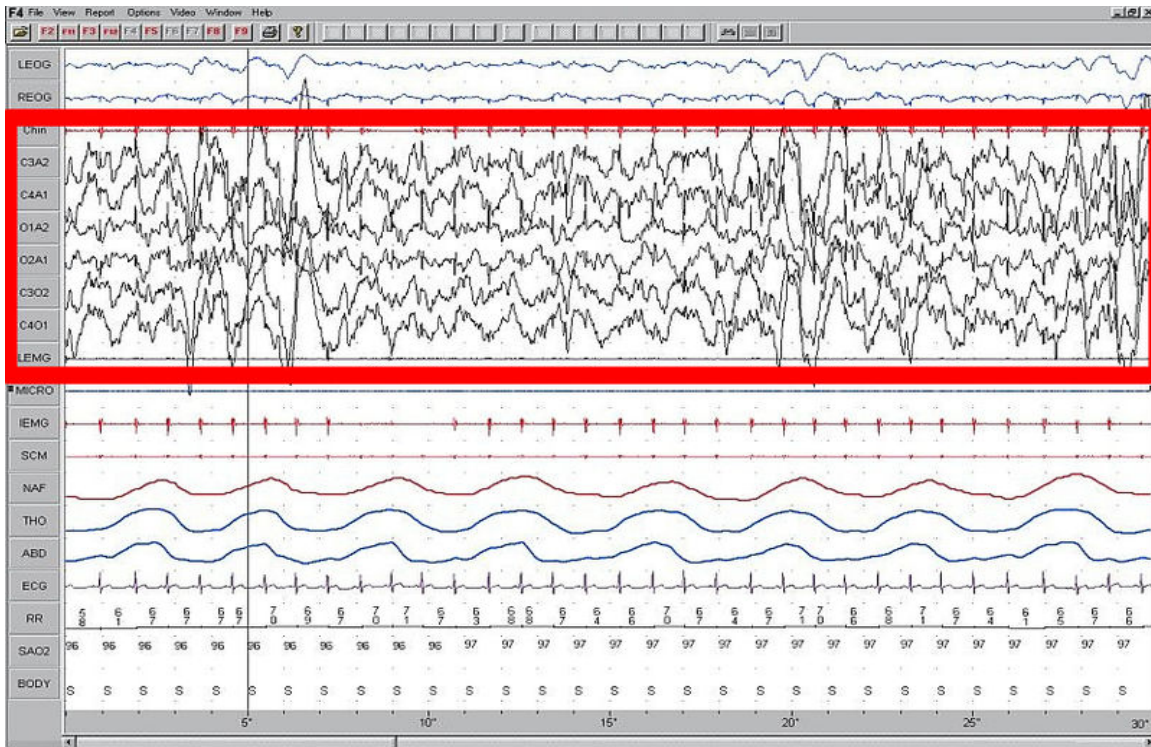
Stage N3 sleep; EEG highlighted by red box. Thirty seconds of deep sleep, here with greater than 50% delta waves.
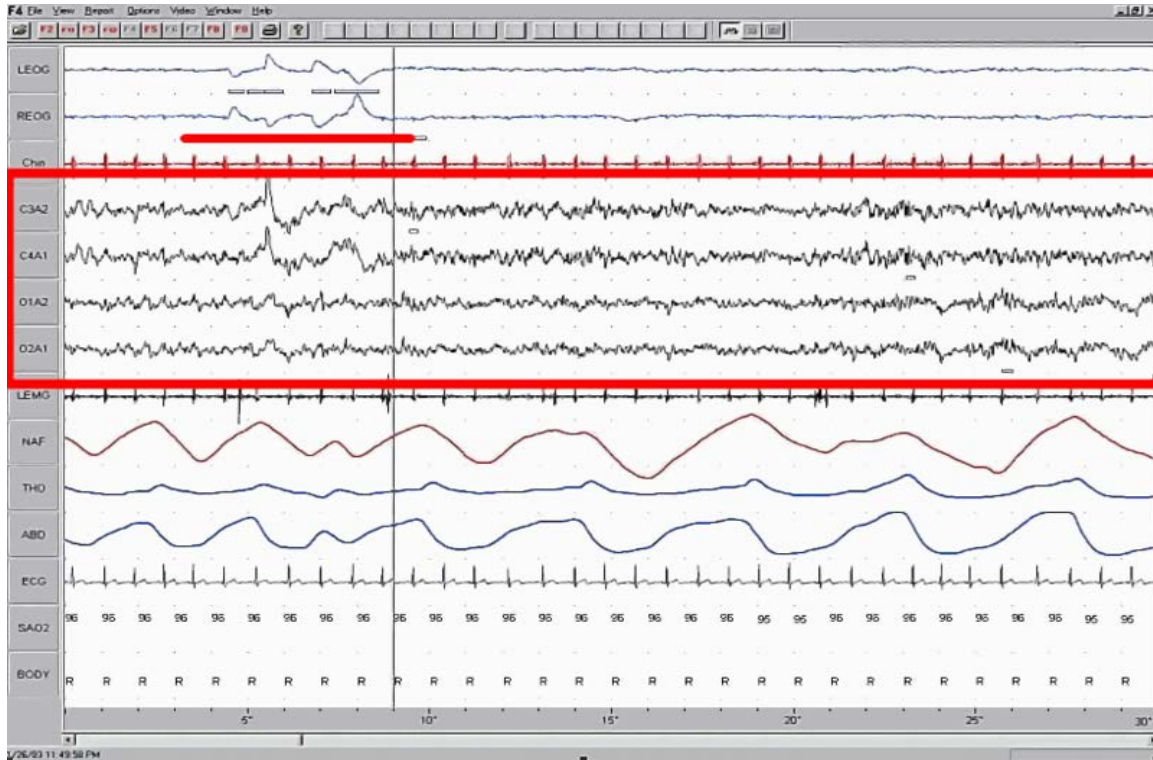
REM sleep; EEG highlighted by red box; eye movements highlighted by red line. Thirty seconds of sleep.

Sleep proceeds in cycles of REM and NREM, the order normally being N1 → N2 → N3 → N2 → REM. There is a greater amount of deep sleep (stage N3) earlier in the sleep cycle, while the proportion of REM sleep increases later in the sleep cycle and just before natural awakening.

The stages of sleep were first described in 1937 by Alfred Lee Loomis and his coworkers, who separated the different electroencephalography (EEG) features of sleep into five levels (A to E), which represented the spectrum from wakefulness to deep sleep. In 1953, REM sleep was discovered as distinct, and thus William Dement and Nathaniel Kleitman reclassified sleep into four NREM stages and REM. The staging criteria were standardized in 1968 by Allan Rechtschaffen and Anthony Kales in the "R&K sleep scoring manual." In the R&K standard, NREM sleep was divided into four stages, with slow-wave sleep comprising stages 3 and 4. In stage 3, delta waves made up less than 50% of the total wave patterns, while they made up more than 50% in stage 4. Furthermore, REM sleep was sometimes referred to as stage 5.

In 2004, the AASM commissioned the AASM Visual Scoring Task Force to review the R&K scoring system. The review resulted in several changes, the most significant being the combination of stages 3 and 4 into Stage N3. The revised scoring was published in 2007 as The AASM Manual for the Scoring of Sleep and Associated Events. Arousals and respiratory, cardiac, and movement events were also added.

Sleep stages and other characteristics of sleep are commonly assessed by polysomnography in a specialized sleep laboratory. Measurements taken include EEG of brain waves, electrooculography (EOG) of eye movements, and electromyography (EMG) of skeletal muscle activity. In humans, each sleep cycle lasts from 90 to 110 minutes on average, and each stage may have a distinct physiological function. This can result in sleep that exhibits loss of consciousness but does not fulfill its physiological functions (i.e., one may still feel tired after apparently sufficient sleep).

Scientific studies on sleep having shown that sleep stage at awakening is an important factor in amplifying sleep inertia. Alarm clocks involving sleep stage monitoring appeared on the market in 2005. Using sensing technologies such as EEG electrodes or accelerometers, these alarm clocks are supposed to wake people only from light sleep.

NREM sleep

According to the 2007 AASM standards, NREM consists of three stages. There is relatively little dreaming in NREM.

Stage N1 refers to the transition of the brain from alpha waves having a frequency of 8–13 Hz (common in the awake state) to theta waves having a frequency of 4–7 Hz. This stage is sometimes referred to as somnolence or drowsy sleep. Sudden twitches and hypnic jerks, also known as positive myoclonus, may be associated with the onset of sleep during N1. Some people may also experience hypnagogic hallucinations during this stage. During N1, the subject loses some muscle tone and most conscious awareness of the external environment.

Stage N2 is characterized by sleep spindles ranging from 11 to 16 Hz (most commonly 12–14 Hz) and K-complexes. During this stage, muscular activity as measured by EMG decreases, and conscious awareness of the external environment disappears. This stage occupies 45–55% of total sleep in adults.

Stage N3 (deep or slow-wave sleep) is characterized by the presence of a minimum of 20% delta waves ranging from 0.5–2 Hz and having a peak-to-peak amplitude >75 µV. (EEG standards define delta waves to be from 0 to 4 Hz, but sleep standards in both the original R&K, as well as the new 2007 AASM guidelines have a range of 0.5–2 Hz.) This is the stage in which parasomnias such as night terrors, nocturnal enuresis, sleepwalking, and somniloquy occur. Many illustrations and descriptions still show a stage N3 with 20–50% delta waves and a stage N4 with greater than 50% delta waves; these have been combined as stage N3.

REM sleep

Rapid eye movement sleep, or REM sleep, accounts for 20–25% of total sleep time in most human adults. The criteria for REM sleep include rapid eye movements as well as a rapid low-voltage EEG. Most memorable dreaming occurs in this stage. At least in mammals, a descending muscular atonia is seen. Such paralysis may be necessary to protect organisms from self-damage through physically acting out scenes from the often-vivid dreams that occur during this stage.

The human biological clock
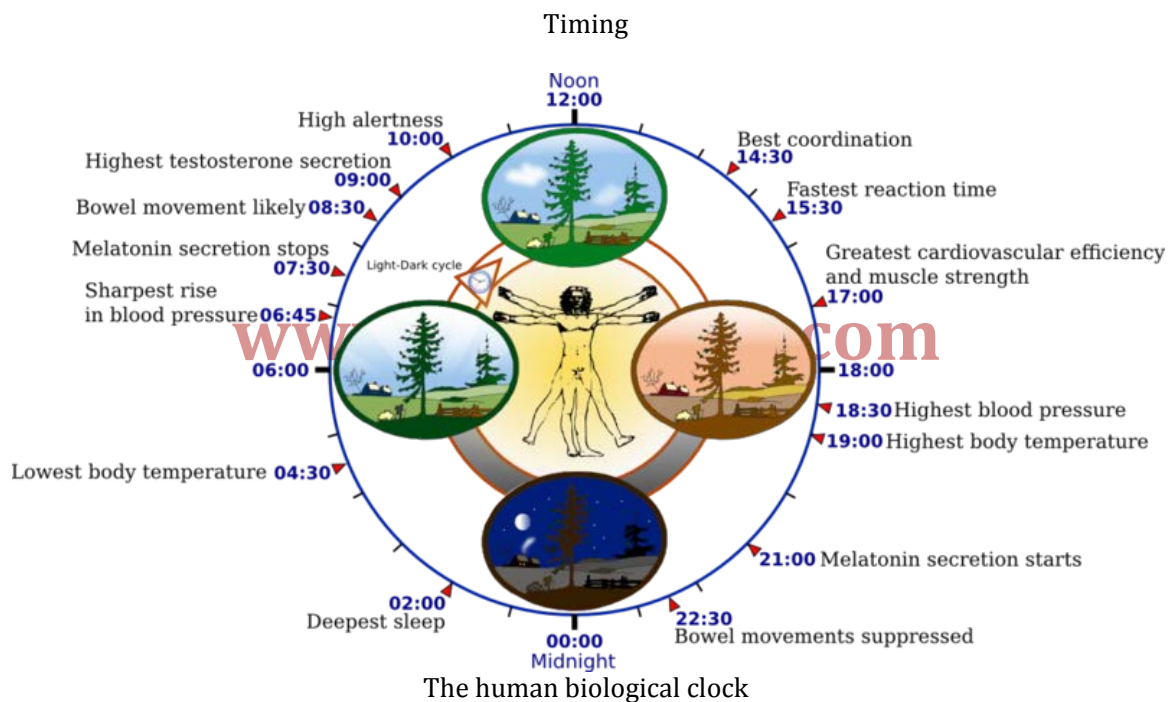
Sleep timing is controlled by the circadian clock, sleep-wake homeostasis, and in humans, within certain bounds, willed behavior. The circadian clock—an inner timekeeping, temperature-fluctuating, enzyme-controlling device—works in tandem with adenosine, a neurotransmitter that inhibits many of the bodily

processes associated with wakefulness. Adenosine is created over the course of the day; high levels of adenosine lead to sleepiness. In diurnal animals, sleepiness occurs as the circadian element causes the release of the hormone melatonin and a gradual decrease in core body temperature. The timing is affected by one's chronotype. It is the circadian rhythm that determines the ideal timing of a correctly structured and restorative sleep episode.

Homeostatic sleep propensity (the need for sleep as a function of the amount of time elapsed since the last adequate sleep episode) must be balanced against the circadian element for satisfactory sleep. Along with corresponding messages from the circadian clock, this tells the body it needs to sleep. Sleep offset (awakening) is primarily determined by circadian rhythm. A person who regularly awakens at an early hour will generally not be able to sleep much later than his or her normal waking time, even if moderately sleep-deprived.

Sleep duration is affected by the gene DEC2. Some people have a mutation of this gene; they sleep two hours less than normal. Neurology professor Ying-Hui Fu and her colleagues bred mice that carried the DEC2 mutation and slept less than normal mice.

Sleep Deprivation

Sleep deprivation is the condition of not having enough sleep; it can be either chronic or acute. A chronic sleep-restricted state can cause fatigue, daytime sleepiness, clumsiness and weight loss or weight gain. It adversely affects the brain and cognitive function. Few studies have compared the effects of acute total sleep deprivation and chronic partial sleep restriction. Complete absence of sleep over long periods is impossible for humans to achieve (unless they suffer from fatal familial insomnia); brief microsleeps cannot be avoided. Long-term total sleep deprivation has caused death in lab animals.

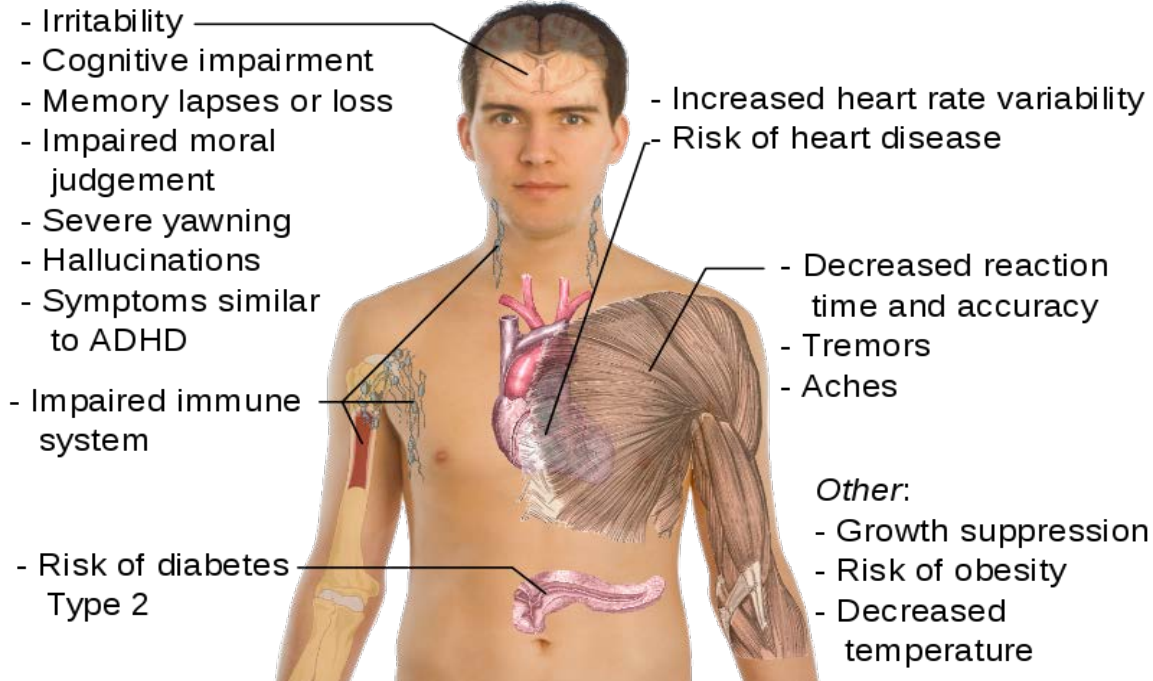### Optimal amount in humans

Adult

The optimal amount of sleep is not a meaningful concept unless the timing of that sleep is seen in relation to an individual's circadian rhythms. A person's major sleep episode is relatively inefficient and inadequate when it occurs at the "wrong" time of day; one should be asleep at least six hours before the lowest body temperature. The timing is correct when the following two circadian markers occur after the middle of the sleep episode and before awakening:

maximum concentration of the hormone melatonin, and
minimum core body temperature.

Human sleep needs can vary by age and among individuals, and sleep is considered to be adequate when there is no daytime sleepiness or dysfunction. Moreover, self-reported sleep duration is only moderately correlated with actual sleep time as measured by actigraphy, and those affected with sleep state misperception may typically report having slept only four hours despite having slept a full eight hours.

A University of California, San Diego psychiatry study of more than one million adults found that people who live the longest self-report sleeping for six to seven hours each night. Another study of sleep duration and mortality risk in women showed similar results. Other studies show that "sleeping more than 7 to 8 hours per day has been consistently associated with increased mortality," though this study suggests the cause is probably other factors such as depression and socioeconomic status, which would correlate statistically. It has been suggested that the correlation between lower sleep hours and reduced morbidity only occurs with those who wake after less sleep naturally, rather than those who use an alarm.

Effects of Sleep deprivation

- Irritability
- Cognitive impairment
- Memory lapses or loss
- Impaired moral judgement
- Severe yawning
- Hallucinations
- Symptoms similar to ADHD

- Impaired immune system

- Risk of diabetes Type 2

- Increased heart rate variability
- Risk of heart disease

- Decreased reaction time and accuracy
- Tremors
- Aches

Other:
- Growth suppression
- Risk of obesity
- Decreased temperature

Main health effects of sleep deprivation, indicating impairment of normal maintenance by sleep

Researchers at the University of Warwick and University College London have found that lack of sleep can more than double the risk of death from cardiovascular disease, but that too much sleep can also be associated with a doubling of the risk of death, though not primarily from cardiovascular disease. Professor Francesco Cappuccio said, "Short sleep has been shown to be a risk factor for weight gain, hypertension, and Type 2 diabetes, sometimes leading to mortality; but in contrast to the short sleep-mortality association, it appears that no potential mechanisms by which long sleep could be associated with increased mortality have yet been investigated. Some candidate causes for this include depression, low socioeconomic status, and cancer-related fatigue... In terms of prevention, our findings indicate that consistently sleeping around seven hours per night is optimal for health, and a sustained reduction may predispose to ill health."

Furthermore, sleep difficulties are closely associated with psychiatric disorders such as depression, alcoholism, and bipolar disorder. Up to 90% of adults with depression are found to have sleep difficulties. Dysregulation found on EEG includes disturbances in sleep continuity, decreased delta sleep and altered REM patterns with regard to latency, distribution across the night and density of eye movements.

Effects on the brain

Sleep deprivation can adversely affect the brain and cognitive function. A 2000 study, by the UCSD School of Medicine and the Veterans Affairs Healthcare System in San Diego, used functional magnetic resonance imaging (fMRI) technology to monitor activity in the brains of sleep-deprived subjects performing simple verbal learning tasks. The study showed that regions of the brain's prefrontal cortex displayed more activity in sleepier subjects. Depending on the task at hand, the brain would sometimes attempt to compensate for the adverse effects caused by sleep deprivation.

The temporal lobe, which is a brain region involved in language processing, was activated during verbal learning in rested subjects but not in sleep-deprived subjects. The parietal lobe, not activated in rested subjects during the verbal exercise, was more active when the subjects were deprived of sleep. Although memory performance was less efficient with sleep deprivation, greater activity in the parietal region was associated with better short term memory.

A 2001 study at Chicago Medical Institute suggested that sleep deprivation may be linked to serious diseases, such as heart disease and mental illness including psychosis and bipolar disorder. The link between sleep deprivation and psychosis was further documented in 2007 through a study at Harvard Medical School and the University of California at Berkeley. The study revealed, using MRI scans, that sleep deprivation causes the brain to become incapable of putting an emotional event into the proper perspective and incapable of making a controlled, suitable response to the event. Sleep deprivation may have been the underlying cause of the overdose deaths of celebrities Heath Ledger and Anna Nicole Smith.

A study tested 17 right-handed civilian males, between the ages of 21–29 years (mean 24.7 ± 2.8 years), with no history of medical, neurological, psychiatric, or sleep disorder conditions. Their histories also included 7–8 h of nightly sleep on a regular basis, no nicotine use, and low caffeine use (less than 100 mg/day). The negative effects of sleep deprivation on alertness and cognitive performance suggest decreases in brain activity and function, primarily in the thalamus, structure involved in alertness and attention, and in the prefrontal cortex, a region sub-serving alertness, attention, and higher-order cognitive processes.

This study used a combination of positron emission tomography (PET) and Fluorine-2-deoxyglucose (FDG), a marker for regional cerebral metabolic rate for glucose (CMRglu) and neuronal synaptic activity. A time series design was used, with progressive sleep deprivation as the independent variable. Repeated measures of absolute regional CMRglu, cognitive performance, alertness, mood, and subjective experiences were collected after 0, 24, 48, and 72 h of sleep deprivation. Additional measures of alertness, cognitive performance, and mood were collected at fixed intervals throughout the sleep deprivation period. These measures were included to place the performance results associated with the PET scans in the context of the circadian rhythm of cognitive performance, as well as to impose a moderate-to-heavy near continuous workload on the subjects as might be anticipated in a real-world sustained operation.

A noted 2002 University of California animal study indicated that non-rapid eye movement sleep (NREM) is necessary for turning off neurotransmitters and allowing their receptors to "rest" and regain sensitivity which allows monoamines (norepinephrine, serotonin and histamine) to be effective at naturally produced levels. This leads to improved regulation of mood and increased learning ability. The study also found that rapid eye movement sleep (REM) deprivation may alleviate clinical depression because it mimics selective serotonin reuptake inhibitors (SSRIs). This is because the natural decrease in monoamines during REM is not allowed to occur, which causes the concentration of neurotransmitters in the brain, that are depleted in clinically depressed persons, to increase. Sleep outside of the REM phase may allow enzymes to repair brain cell damage caused by free radicals. High metabolic activity while awake damages the enzymes themselves preventing efficient repair. This study observed the first evidence of brain damage in rats as a direct result of sleep deprivation.

Animal studies suggest that sleep deprivation increases stress hormones, which may reduce new cell production in adult brains.

Attention and working memory

Among the numerous physical consequences of sleep deprivation, deficits in attention and working memory are perhaps the most important; such lapses in mundane routines can lead to unfortunate results, from forgetting ingredients while cooking to missing a sentence while taking notes. Working memory is tested by such methods as choice-reaction time tasks.

The attentional lapses also extend into more critical domains in which the consequences can be literally life-or-death; car crashes and industrial disasters can result from inattentiveness attributable to sleep deprivation. To empirically measure the magnitude of attention deficits, researchers typically employ the psychomotor vigilance task (PVT) which requires the subject to press a button in response to a light at pseudo-random intervals. Failure to press the button in response to the stimulus (light) is recorded as an error, attributable to the microsleeps that occur as a product of sleep deprivation.

Crucially, individuals' subjective evaluations of their fatigue often do not predict actual performance on the PVT. While totally sleep-deprived individuals are usually aware of the degree of their impairment, lapses from chronic (lesser) sleep deprivation can build up over time so that they are equal in number and severity to the lapses occurring from total (acute) sleep deprivation. Chronically sleep-deprived people, however, continue to rate themselves considerably less impaired than totally sleep-deprived participants. Since people usually evaluate their capability on tasks like driving subjectively, their evaluations may lead them to the false conclusion that they are able to perform tasks that require constant attention when their abilities are in fact impaired.

Impairment of ability

The dangers of sleep deprivation are apparent on the road; the American Academy of Sleep Medicine (AASM) reports that one in every five serious motor vehicle injuries is related to driver fatigue, with 80,000 drivers falling asleep behind the wheel every day and 250,000 accidents every year related to sleep, though the National Highway Traffic Safety Administration suggests the figure for traffic accidents may be closer to 100,000. The AASM recommends pulling off the road and taking a 15- or 20-minute nap to alleviate drowsiness.

According to a 2000 study published in the British Medical Journal, researchers in Australia and New Zealand reported that sleep deprivation can have some of the same hazardous effects as being drunk. People who drove after being awake for 17–19 hours performed worse than those with a blood alcohol level of .05 percent, which is the legal limit for drunk driving in most western European countries and Australia. Another study suggested that performance begins to degrade after 16 hours awake, and 21 hours awake was equivalent to a blood alcohol content of .08 percent, which is the blood alcohol limit for drunk driving in Canada, the U.S., and the U.K.

In addition, as a result of continuous muscular activity without proper rest time, effects such as cramping are much more frequent in sleep-deprived individuals. Extreme cases of sleep deprivation have been reported to be associated with hernias, muscle fascia tears, and other such problems commonly associated with physical overexertion.

A 2006 study has shown that while total sleep deprivation for one night caused many errors, the errors were not significant until after the second night of total sleep deprivation. However, combining alcohol with acute sleep deprivation results in a trebled rate of driving off the road when using a simulator.

The National Sleep Foundation identifies several warning signs that a driver is dangerously fatigued, including rolling down the window, turning up the radio, trouble keeping eyes open, head-nodding, drifting out of the lane, and daydreaming. At particular risk are lone drivers between midnight and 6 a.m.

Sleep deprivation can negatively impact performance in professional fields as well, potentially jeopardizing lives. Due largely to the February 2009 crash of a regional jet in Buffalo, NY, which killed 50 people and was partially attributed to pilot fatigue, which caused the FAA to review its procedures to ensure pilots are sufficiently rested. A 2004 study also found medical residents with less than four hours of sleep a night made more than twice as many errors as residents who slept for more than seven hours a night, an especially alarming trend given that less than 11% of surveyed residents were sleeping more than seven hours a night.

Twenty-four hours of continuous sleep deprivation results in the choice of less difficult math tasks without decreases in subjective reports of effort applied to the task. Naturally caused sleep loss affects the choice of everyday tasks such that low effort tasks are mostly commonly selected. Adolescents who experience less sleep show a decreased willingness to engage in sports activities that require effort through fine motor coordination and attention to details.

Great sleep deprivation mimics psychosis: distorted perceptions can lead to inappropriate emotional and behavioral responses.

Astronauts have reported performance errors and decreased cognitive ability during periods of extended working hours and wakefulness as well as due to sleep loss caused by circadian rhythm disruption and environmental factors.

(c) Enumerate various methods of problem solving. Critically evaluate their advantages and limitations also.

TOPIC: *Thinking and Problem Solving*
SUBTOPIC: *Reasoning and problem solving*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

PROBLEM SOLVING

Problem solving is defined as the goal-driven process of overcoming obstacles that obstruct the path to a solution (Simon, 1999a; Sternberg, 1999). Problem solving and reasoning are alike in many ways. For example, in both problem solving and reasoning, the individual is creating new knowledge, albeit in the form of a solution needed to reach a goal or in the form of a conclusion derived from evidence, respectively. Problem solving and reasoning seem to differ, however, in the processes by which this new knowledge is created. In problem solving, individuals use strategies to overcome obstacles in pursuit of a solution (Newell & Simon, 1972). In reasoning, however, the role of strategies is not as clear. It was mentioned earlier that reasoning theories, such as syntactic rule theory, pragmatic reasoning theory, and mental model theory, do not explicitly specify if syntactic rules, pragmatic reasoning schemas, and mental models, respectively, should be viewed as strategies or, more fundamentally, as forms of *representing* knowledge. *Representation* refers to the way in which knowledge or information is formalized in the mind, whereas *strategy* refers to the methods by which this knowledge or information is manipulated to reach a goal. Although individuals may be consciously aware of the strategies they choose to solve problems, individuals are believed to be unaware of how they represent knowledge, which is considered to be part of the mind's cognitive architecture.

It is possible that strategies are unimportant in reasoning because the objective in reasoning is not to reach a goal so much as it is to infer what follows from evidence; the conclusion is meant perhaps to fall out of the set of premises without too much work on the part of the reasoner. Although some reasoning tasks do require goal-oriented conclusions that are not easily deduced—or directly deduced at all—from the premises, it might be more accurate to describe such reasoning tasks as more akin to problem-solving tasks (Galotti, 1989; Evans, Over, & Manktelow, 1993). For instance, reasoning tasks leading to inductive inferences—inferences that go beyond the information given in the task—might be considered more akin to problem-solving tasks. Strategies, however, are clearly important in problem solving because the goal in problem solving is to reach a solution, which is not always derived deductively or even solely from the problem information.

Knowledge Representation and Strategies in Problem Solving

Production Systems

The distinction between representation and strategy is made explicit in the problem-solving literature. For example, some investigators propose that knowledge is represented in terms of *production systems* (Dawson, 1998; Simon, 1999b; Sternberg, 1999). In a production system, instructions (called productions) for behavior take the following form:

IF<<conditions> , THEN<<actions>.

The form above indicates that if certain conditions are met or satisfied, then certain actions can be carried out (Simon, 1999b). The conditions of a production involve propositions that "state properties of, or

relations among, the components of the system being modeled" (Simon, 1999b, p. 676). A production system is normally implemented following a match between the conditions of the production and elements stored in working memory. The production is implemented when the conditions specified in the production's IF clause are satisfied or met by the elements of working memory. Following the satisfaction of the production system's IF clause with the elements of working memory, an action is initiated (as specified in the production system). The action may take the form of a motor action or a mental action such as the elimination or creation of working memory elements (Simon, 1999b).

The elements of working memory may satisfy the conditions of numerous productions at any given time. One way in which all the productions that are executable at a given moment can be restrained from overwhelming the problem solver is through the presence of goals. A *goal* can be defined simply as a symbol or representation that must be present both in the conditions of the production and in working memory before that production is activated. In other words, a goal provides a more stringent condition that must be met by an element in working memory before the production is activated (Simon, 1999b). In the following example of a production system, the goal is to determine if a particular sense of the word *knows* is to be applied (taken from Lehman, Lewis, & Newell, 1998, p. 156):

IF comprehending *knows,* and
there's a preceding word, and
that word can receive the subject role, and
the word refers to a person, and
the word is third person singular,
THEN use sense 1 of *knows.*

The antecedent or the condition of the production consists of a statement of the goal (i.e., comprehending *knows*), along with additional conditions that need to be met before the consequent or action is applied (i.e., use sense 1 of *knows*). Although the above production system might look like a strategy, it is not because knowledge has not been manipulated.

Parallel Distributed Processing (PDP) Systems

Other theories of knowledge representation exist outside of production systems. For example, some investigators propose that knowledge is represented in the form of a parallel distributed processing (PDP) system (Bechtel & Abrahamsen, 1991; Dawson, 1998; Dawson, Medler, & Berkeley, 1997). A PDP system involves a network of inter-connected, processing units that learn to classify patterns by attending to their specific features. A PDP system is made up of simple processing units that communicate information about patterns by means of weighted connections. The weighted connections inform the recipient processing unit whether a to-be classified pattern includes a feature that the recipient processing unit needs to attend to and use in classifying the pattern. According to PDP theory, knowledge is represented in the layout of connections that develops as the system learns to classify a set of patterns. In Figure 23.3, a PDP representation of the Wason (1966) selection task is shown. This representation illustrates a network that has learned to select the *P* and *Q* in response to the selection task (Leighton & Dawson, 2001). The conditional rule and set of four cards are coded as *1*s and *0*s and are presented to the network's input unit layer. The network responds to the task by turning on one
of the four units in its output unit layer, which correspond to the set of four cards coded in the input unit layer. The layer of hidden units indicates the number of cuts or divisions in the pattern space required to solve the task correctly (i.e., generate the correct responses to the task). Training the network to generate the *P* response required a minimum of three hidden units.
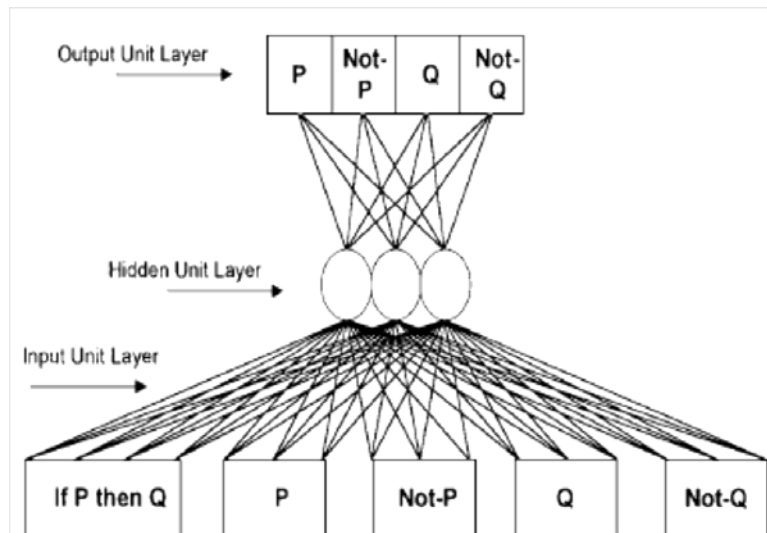
Figure 23.3 Illustration of a PDP network, including layer of input units, hidden units, and output units (adapted from Leighton & Dawson, 2001).

Strategies can be extracted from a PDP system. The process by which strategies are identified in a PDP system is laborious, however, and requires the investigator to examine the specific procedures used by the system to classify a set of patterns (Dawson, 1998).

Algorithms

The representation of knowledge provides the language in which cognitive processes in models of cognitive systems can be described. An *algorithm* is one cognitive process for accomplishing an explicit outcome. More specifically, an algorithm is made up of a finite set of operations that is straightforward and unambiguous and, when applied to a set of objects (e.g., playing cards, chess pieces, computer parts), leads to a specified outcome (Dietrich, 1999). The initial state of the set of objects constitutes the input to the algorithm, and the final state of the objects constitutes the output of the algorithm. The initial state of objects is transformed into a final state by implementing the operations of the algorithm that correspond to state transitions. Algorithms can be described more specifically when the context of the algorithm is defined because an algorithm's clarity and simplicity are relative to the context in which it is being applied (Dietrich, 1999). An example of an algorithm might be the instructions included with a new desktop computer (at least, such instructions are supposed to be algorithms). If one follows the instructions for installing all the parts of the computer, the outcome is certain: a working computer. Algorithms are sometimes unavailable for accomplishing certain outcomes; under these circumstances, heuristics can be implemented to *approximate* the desired outcome.

Heuristics

A problem-solving *heuristic* is a rule of thumb for approximating a desired outcome. As with reasoning heuristics, problem-solving heuristics sometimes produce desired outcomes and sometimes not. Heuristics are imperfect strategies (Fischhoff, 1999). Examples of heuristics are considered below in the context of Newell and Simon's model of problem solving.

Theories of Problem Solving

Newell and Simon's Model of Problem Solving

Even after 25 years, Newell and Simon's (1972) model of problem solving remains influential today. Newell and Simon's model of problem solving was generated from computer simulations and from participants' think-aloud responses as they worked through problems. According to the model, the problem solver perceives both the *initial state,* the state at which he or she originally is, and the *goal state,* the state that the

69

problem solver would like to achieve. Both of these states occupy positions within a *problem space,* the universe of all possible actions that can be applied to the problem, given any constraints that apply to the solution of the problem (Simon, 1999a; Sternberg, 1999).

In the ongoing process of problem solving, a person decomposes a problem into a series of intermediate steps with the purpose of bringing the initial state of the problem closer to the goal state. At each intermediate step prior to the goal state, the sub-goal is to achieve the next intermediate step that will bring the problem solver closer to the goal state. Each step toward the goal state involves applying an operation or rule that will change one state into another state. The set of operations is organized into a program, including sublevel programs. The program can be a heuristic or an algorithm, depending on its specific nature. In short, according to Newell and Simon's (1972) model, problem solving is a search through a series of states within a problem space; the solution to a problem lies in finding the correct sequence of actions for moving from one (initial) state to another (goal) state (Newell & Simon, 1972; Simon, 1999a; Sternberg, 1999).

A variety of heuristics can be used for changing one state into another. For example, the *difference-reduction method* involves reducing the difference between the initial state and goal state by applying operators that increase the surface similarity of both states. If an operator cannot be directly applied to reduce the difference between the initial state and goal state, then the heuristic is discarded. Another method that is similar to the difference-reduction method is Newell and Simon's (1972) *means-ends analysis,* a heuristic Newell and Simon studied extensively in a computer simulation program (i.e., General Problem Solver [GPS]) that modeled human problem solving. Means-end analysis is similar to the difference-reduction method, with the exception that if an operator cannot be directly applied to reduce a difference between the initial state and goal state, then, instead of the strategy's being discarded, a sub-goal is set up to make the operator applicable (Simon, 1999a).

Analogy is another heuristic. Under this heuristic, the problem solver uses the structure of the solution to an analogous problem to guide his or her solution to a current problem. The main focus in research on analogy is in how people interpret or understand one situation in terms of another; that is, how it is that one situation is *mapped* onto another for problem-solving purposes (Gentner, 1999). Two main sub-processes are proposed to mediate the use of analogy. According to Gentner's structure-mapping theory (1983), an unfamiliar situation can be understood in terms of another familiar situation by aligning the representational structures of the two situations and projecting inferences from the familiar case to the unfamiliar case. The alignment must be structurally consistent such that there is a one-to-one correspondence between the mapped elements in the familiar and unfamiliar situations. Inferences are then projected from the familiar to the unfamiliar situation so as to obtain structural completion (Gentner, 1983, 1999). Following this alignment, the analogy and its inferences are *evaluated* by assessing (a) the structural soundness of the alignment between the two situations; (b) the factual validity of the inferences, because the use of analogy does not guarantee deductive validity; and (c) whether the inferences meet the requirements of the goal that prompted the use of the analogy in the first place (Gentner, 1999).

Recent research suggests that use of analogy in real-world contexts is based on structural or deep underlying similarities, instead of surface or superficial similarities, between the unfamiliar situation and the familiar situation (Dunbar, 1995, 1997). For example, Dunbar (1997) found that over 50% of analogies that scientists generated at weekly meetings in a molecular biology lab were based on deep, structural features between problems, rather than on surface features between problems. In previous studies, however, investigators (e.g., Gentner, Rattermann, & Forbus, 1993) have found that participants in laboratory experiments sometimes rely on superficial features when using analogy. According to Blanchette and Dunbar (2000; see also Dunbar, 1995, 1997), participants' reliance on superficial features when using analogy might be due to the kind of paradigm used to study analogy. For example, Blanchette and Dunbar indicated that previous studies have used a *reception paradigm* to study analogy use. Under the reception paradigm, participants are provided with both a target (less familiar) and a source (familiar) analog and then asked to indicate the relationships between both rather than being asked to generate their own source analogs. In a series of studies aimed at evaluating participants' analogies, Blanchette and Dunbar found that when participants were given a target problem and asked to *generate* their own source analog, most of the analogies (67%) generated by participants did not exhibit superficial similarities with

the target but, instead, exhibited deeper similarities with the target. The proportion of these deep analogies increased to 81% when participants worked individually. These results suggest that participants, like scientists, can generate analogies based on deep, structural features when laboratory conditions are more akin to real-world contexts, that is, when participants are free to generate their own source analogs.

Error is always a possibility when heuristics are used. Not only might a chosen heuristic be inappropriate for the problem under consideration, but a heuristic might be inappropriately used, resulting in unsuccessful problem solving. Heuristics such as the difference-reduction method, means-end analysis, analogy, and others (e.g., see Anderson, 1990, for further descriptions of the generate and test method, working forward method, and working backward method) are only general rules of thumb that work most of the time but not necessarily all of the time (Fischhoff, 1999; Holyoak, 1990; Simon, 1999a). They represent general problem-solving methods that can be applied with relative success to a wide range of problems across domains.

According to Newell and Simon (1972), the use of heuristics embodies problem solving because of the cognitive limitations or *bounded rationality* that characterizes human behavior (see also Sternberg & Ben Zeev, 2001). Simon (1991) described bounded rationality as involving two central components: the limitations of the human mind and the structure of the environment in which the mind must operate. The first of these components suggests that the human mind is subject to limitations, and, due to these limitations, models of human problem solving, decision making, and reasoning should be constructed around how the mind actually performs instead of on how the mind should perform from an engineering point of view. Foolproof strategies do not exist in everyday cognition because the ill-defined structure of our environment makes it unlikely that people can identify perfect heuristics for solving imperfect, uncertain problems. The second of these components suggests that the structure of the environment shapes the heuristics that will be most successfully applied in problem solving endeavors. If the environment is ill defined (in the sense that it reflects numerous uncertain tasks), then general heuristics that work most of the time and do not overburden the cognitive system will be favored (see also Brunswick, 1943; Gigerenzer et al., 1999; Shepard, 1990). Heuristics, however, are only one of the kinds of tools that facilitate problem solving. Investigators have also found that *insight* is an important variable that aids some forms of problem solving (Davidson & Sternberg, 1984; Metcalfe & Wiebe, 1987; Sternberg & Davidson, 1995).

Problem Solving by Means of Insight

Insightful problem solving can be defined as problem solving that is significantly assisted by the awareness of a key piece of information—information that is not necessarily obvious from the problem presented (Sternberg, 1999). It is believed that insight plays a role in the solution of *ill-defined problems*. Ill-defined problems are problems whose solution paths are elusive; the goal is not immediately certain. Because the solution path is elusive, ill-defined problems are challenging to represent within a problem space. Ill-defined problems are often termed *insight problems* because they require the problem solver to perceive the problem in a new way, a way that illuminates the goal state and the path that leads to a solution. Insight into a solution can manifest itself after the problem solver has put the problem aside for hours and then comes back to it. The new perspective one gains on a problem when coming back to it after having put it aside is known as an *incubation effect* (Dominowski & Jenrick, 1973; Smith & Blankenship, 1989).

Metcalfe and Wiebe (1987; see also Metcalfe, 1986, 1998) have shown that insightful problem solving seems to differ from ordinary (non-insightful) problem solving. For example, these investigators have shown that participants who are highly accurate in estimating their problem-solving success with ordinary problems are not as accurate in estimating their success with insight problems. The processes that might be responsible for these differences are not yet detailed, making this account more representative of a performance-based account than a process-based account of problem solving (for a fuller discussion of insight, see Sternberg & Davidson, 1995).

In a more process-oriented theory of insight, however, Davidson and Sternberg (1984) have offered a three process view of insight. These investigators have proposed that insightful problem solving manifests itself in three different forms: (a) *Selective encoding* insights involve attending to a part of the problem that is

71

©VISION IAS
www.visionias.wordpress.com
Contact: +91-9650617807, +91-9968029039

©Numerons
www.numerons.com
Contact: +91-9560344245

relevant to solving the problem, (b) *selective comparison* insights involve novel comparisons of information presented in the problem with information stored in long-term memory, and (c) *selective combination* insights involve new ways of integrating and synthesizing new and old information. Insight gained in any one of these three forms can facilitate insightful problem solving.
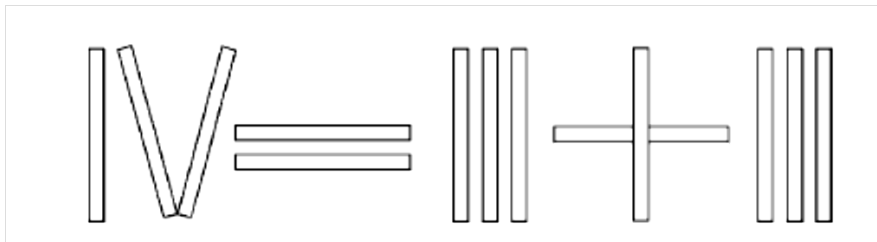
Figure 23.4 Example of matchstick problem (adapted from Knoblich, Ohlsson, Haider, & Rhenius, 1999).

In addition, Knoblich, Ohlsson, Haider, and Rhenius (1999) have characterized insightful problem solving as overcoming *impasses,* states of mind in which the thinker is unsure of what to do next. These investigators have proposed that impasses are overcome by changing the problem representation by means of two hypothetical processes or mechanisms. The first mechanism involves *relaxing* the constraints imposed upon the solution, and the second mechanism involves decomposing the problem into *perceptual chunks*. In a series of four studies aimed at examining insightful problem solving, Knoblich et al. (1999) asked participants to solve insight problems called "match-stick arithmetic" problems. As shown in Figure 23.4, match-stick arithmetic problems involve false arithmetic statements written with Roman numerals (e.g., I, II, IV), arithmetic operations (e.g., −, +), and equal signs constructed out of matchsticks. The goal in matchstick problems is to move a single stick in such a way that the initial false arithmetic statement is transformed into a true statement. A move can be made on a numerical value or an operator and can consist of grasping a stick and moving it, rotating it, or sliding it.

According to Knoblich et al. (1999), matchstick problems can be solved by relaxing the constraints on how numerical values are represented, how operators are represented, and how arithmetic functions are supposed to be formed—for example, form of X = f(Y, Z). In particular, the numerical value constraint in arithmetic suggests that a numerical value on one side of an equation cannot be changed unless an equivalent change is made to the numerical value on the other side of the equation, such as when the same quantity is added to or subtracted from both sides of an equation. Relaxing the constraint on how numerical values are represented would involve accepting the possibility that a numerical value on one side of an equation can be changed without changing the other side of the equation as well (e.g., if 1 is subtracted from one side of the equation, this same operation need not be performed on the other side of the equation). Note that numerical value constraints do not include constraints on how the numerical quantities are perceived. For example, the numerical value constraint does not include constraints on whether the number 4 is perceived as *IV* or as *IIII* or some other representation. According to Knoblich et al. (1999), how numbers are perceived in the context of the matchstick task is better explained by considering the process of chunking.

Knoblich et al. (1999) suggest that decomposing elements of matchstick problems into perceptual chunks can also help to solve the problems. Perceptual decomposition involves, for instance, recognizing that the Roman numeral *IV* can be decomposed into the elements *I* and *V,* and that the resulting elements can be moved independently of each other to generate a true matchstick arithmetic equation. Roman numerals cannot, however, be decomposed into elements that are not used in constructing the numerals. For instance, the Roman numeral *IV* could not be decomposed into *IIII* because four vertical lines were not used to construct the numeral *IV*.

In an effort to examine how constraint relaxation and chunking mediated insightful problem solving, Knoblich et al. (1999) asked participants to solve matchstick problems of varying difficulty. After an initial training phase, participants were presented with two blocks of six matchstick problems on a computer screen and given 5 minutes to respond to each problem. Each block of problems contained instances of easy matchstick problems (i.e., Type A) and difficult matchstick problems (i.e., Type C and D). Results from their four studies revealed, as expected, that participants were more successful at solving problems that required

72

the relaxation of lower order constraints (e.g., relaxing constraints on numerical value representation) than problems that required the relaxation of higher order constraints (e.g., relaxing constraints on arithmetic function representation). For example, after an average of 5 minutes, almost all participants solved problems requiring the relaxation of low-order constraints (Type A), whereas fewer than half of all participants solved problems requiring the relaxation of high-order constraints (Type C). In addition, participants were more successful at solving problems that required the decomposition of loose chunks (e.g., decomposing *IV* into *I* and *V*) than problems that required the decomposition of tight chunks (e.g., decomposing *V* into \ and /). After an average of 5 minutes, almost all participants solved problems requiring the decomposition of loose chunks (Type A), whereas only 75% of participants solved problems requiring the decomposition of tight chunks (Type D). Overcoming impasses in solving insight problems exemplifies a general need to override *mental sets* or fixed ways of thinking about problems generated from past experience with similar problems. The encumbrance of mental sets highlights the existence of factors such as how the problem is interpreted that can influence problem-solving success.

It is very likely that Oedipus solved the sphinx's riddle by experiencing an insight into its solution. The riddle can certainly be labeled an ill-defined problem—one whose solution required the awareness of a key piece of information. What are the processes by which Oedipus gained the insight necessary to solve the riddle? This is an important question, but one whose answer remains a mystery. On the one hand, that any belief or thought can, in principle, be brought to bear on problem-solving endeavors permits the possibility of creative or insightful problem solving. On the other hand, because any belief or thought can be brought to bear on problem-solving endeavors, understanding how individuals select specific beliefs and thoughts as they solve problems remains a challenge—a challenge that we earlier identified as the frame problem (Fodor, 1983).

Factors that Mediate Problem Solving

Definition of Problem: Mental Set

A mental set involves thinking about a problem, its context, and its possible solution from a single perspective (Luchins, 1942; Sternberg, 1999). Such a limited perspective can hinder problem solving if a successful solution can be achieved only by viewing the problem from a novel angle. Setting the problem aside momentarily can foster insight or a new perspective (see earlier discussion of *incubation effect*) and help break the mental set. For example, misreading a word in an essay or misreading a variable in a mathematical proof can lead to a mental set and block understanding. In these cases, putting the material aside even for an hour and then coming back to it can break the mental set.

Past experience can be beneficial to problem solving, but it can also foster mental sets by biasing the way in which the problem solver ventures to reach a solution. In particular, *expertise* in the domain of the problem can actually disrupt problem solving, especially if the problem calls for a creative solution (Wiley, 1998). Although experts are generally able to solve problems in their domains more effectively than novices because their well-structured, easily activated knowledge permits an efficient search of the problem space, sometimes this knowledge can be disadvantageous. For example, Wiley (1998) has suggested that a large amount of domain knowledge can bias problem-solving efforts by confining the search space and therefore excluding the portion of the space in which the solution resides. That is, expertise can actually constrain creative problem solving by foreclosing the problem space prematurely (see also Bedard & Chi, 1992; Frensch & Sternberg, 1989).

Strategy Selection and Knowledge

Selecting the right strategy in response to a problem can determine whether a problem's solution will be found and, if so, whether it will be found expeditiously. For example, the *generate and test* heuristic (Newell & Simon, 1972), which involves arbitrarily generating solution paths until the correct path is found, may ultimately lead one down the correct solution path, but it is not a very efficient strategy. In contrast, a *working forward* strategy is more efficient because it involves delimiting the set of possible solution paths and then choosing from this set the one that generates the better solution to the problem. Knowing which

strategy to use in solving a given problem, however, is dependent on the problem solver's level of expertise in the problem domain.

Not all strategies are used equally often by all problem solvers. Strategy selection depends on the problem domain and on the level of expertise of the problem solver within that domain (Chi et al., 1988). Expertise plays a pivotal role in strategy selection because greater domain knowledge in the domain of the problem influences the way in which the problem is interpreted, how the solution is envisioned, and hence the strategy that is ultimately selected to solve the problem. Bedard and Chi (1992), in a review of studies of expert problem solving, concluded that, in general, experts are better problem solvers than are novices because (a) they know more about their domain than do novices; (b) their knowledge is better organized in ways that make that knowledge more accessible, functional, and efficient; (c) they perform better than novices in domain-related tasks on the basis of their greater knowledge and better organization; and (d) their skills are domain specific. In short, experts select strategies and solve problems more efficiently than do novices.

7. Answer the following, each in not more than 400 words: **30x2=60**

(a) What is the essence of a projective technique? Critically evaluate perceptive, apperceptive and productive projective tests in the measurement of personality.

TOPIC:	*Personality*
SUBTOPIC:	*Measurement of Personality (projective tests, pencil-paper test)*
LEVEL:	*Medium*
NATURE:	*Fundamental*

REFERENCE:

Anastasi and Urbina (1996) have characterized a *projective test* as a "relatively unstructured task, that is, a task that permits almost an unlimited variety of possible responses. In order to allow free play to the individual's fantasy, only brief, general instructions are provided" (p. 411). This global, descriptive definition identifies some important elements of projective tests. Ironically, however, this definition and others like it impede our understanding of the nature of projective tests when they are casually juxtaposed with so-called *objective tests*. Without pause, many American psychologists categorize tests according to the traditional projective-objective dichotomy. In thinking and communicating about assessment instruments, these psychologists treat the characteristics of each class of instrument as mutually exclusive or as polar opposites. For example, because objective tests are thought of as unbiased measures, projective tests, by default, are assumed to be subjective. As another example, because objective tests are seen as having standardized administration and scoring, projective tests are assumed to lack empirical rigor. There are a number of reasons that the projective-objective dichotomy leads to an oversimplified and biased understanding of projective tests. First, the projective-objective dichotomy often results in misleading reductionism. Instruments under the rubric of *projective* are assumed to be uniform in content, purpose, and methodology. For example, all projective instruments are often reduced and treated as equivalent to a classic exemplar such as the Rorschach. Reducing all projective instruments to the Rorschach ignores their incredible diversity. Not only do these tests target many different domains of functioning, but they also employ a great variety of methodologies for the purposes of inducing very different response processes. For example, early instruments included an indistinct speech interpretation, word association, cloud perception, hand-positioning perception, comic strip completion, and musical reverie tests (Anastasi & Urbina; Campbell, 1957; Frank, 1939/1962; Murray, 1938). Moreover, this great variety suggests that projective processes are ubiquitous and are involved in many real-life behaviors.

Second, the projective-objective dichotomy implies that there are characteristics unique to each class of test, but these supposed hallmarks are misleading. For example, test elements identified as projective, such as the flexible response format and ambiguous or incomplete stimuli, are employed by tests generally considered to be models of objectivity and quantification. Murstein (1963) notes from the flexible response format of some cognitive ability tests that "we learn a great deal about the person who, on the vocabulary subtests of the Wechsler Adult Scale of Intelligence, when asked to give the meaning of the word 'sentence,'

74

proceeds to rattle off three or four definitions and is beginning to divulge the differences between the connotations and denotations of the word when he is stopped" (p. 3). E. Kaplan's (1991) approach to neuropsychological testing focuses on process, similar to the response-process approach in projective testing. Similarly, Meehl points out the projective element of stimulus ambiguity in self-report personality tests. In his *Basic Readings on the MMPI: A New Selection on Personality Measurement* (1945/1980), Meehl notes that many Minnesota Multiphasic Personality Inventory (MMPI) items, such as "Once in a while I laugh at a dirty joke," contain ambiguities. At the most basic level, it is unclear whether "once in a while" refers to once a day, once a week, or once in a month. Third, the stereotypic juxtaposition of objective and projective testing lends a pejorative connotation to projective tests that suggests they lack objectivity. This is misleading. Many projective tests are quantified and standardized in terms of administration, and more should be. If we take the example of cognitive tests, the style or process of the response can be systematically observed, quantified, and standardized. This qualitative-to-quantitative test development strategy is exactly the same procedure used in sophisticated quantification of projective tests, as in the Rorschach Comprehensive System (Exner, 1993) and the Washington Sentence Completion Test (Loevinger & Wessler, 1970). Such approaches can result in psychometrically sound quantification and standardization. For example, Joy, Fein, Kaplan, and Freedman (2001) utilized this procedure to standardize observation of the Block Design subtest from the Wechsler scales. Other research summarized by Stricker and Gold (1999) and Weiner (1999) indicates that behavioral observation within projective tests can be used to elaborate previously developed hypotheses and to synthesize inferences about the respondent. These same authors also demonstrated these tactics in case examples.

Of course, quantification and reducing examiner bias, that is variability introduced by examiners, are important goals in improving psychological assessment. Nonetheless, reducing examiner variability is not the only goal of assessment and is not equivalent to validity and utility. Indeed, further research should address the extent to which the examiner's input is induced by the subject, as would be the case with reciprocal determinism, increasing the ecological validity of projective tests (Bandura, 1978; Viglione & Perry, 1991). Furthermore, one may speculate that overemphasis on eliminating examiner variability to achieve objectivity can increase test reliability at the expense of validity when it limits salient observations by the examiner.

Finally, projective and objective tests resemble each other in that they share the same goal: the description of personality, psychopathology, and problems in living. However, the dichotomy highlights the differences in method and overlooks fundamental differences in their approach to understanding personality. Later sections of this chapter will highlight some of these differences. As we shall see, the differences may be more in the philosophy of the psychologist using the tests rather than in the tests themselves. The foregoing are only a few examples of the distortions involved in the unexamined use of the projective-objective dichotomy of tests. Furthermore, this familiar dichotomy damages the reputation of projective testing and misleads students. A more informed approach to projective testing is needed. Along those lines, we will juxtapose projective tests against self-report tests in the remainder of this chapter.

**PROBLEMS WITH COMMON METAPHORS AND MODELS**

Like the distinction between projective and objective tests, the common metaphors and models used to describe the projective response process can be grossly misleading. The two well-known metaphors of the projective response process are the blank screen and the X-ray machine. Each metaphor contains an implicit theoretical model of projective testing that shapes our understanding of the projective response process. In this section we critically examine both metaphors.

The Blank Screen Metaphor

The most common and stereotypic metaphor is that of the *blank screen*. In this metaphor, a projective test stimulus is portrayed as a blank screen or canvas upon which the respondent projects his or her inner world (Anastasi & Urbina, 1996). In the reductionistic application of this metaphor, response content is treated as a direct representation of the respondent's inner life. For example, when a respondent projects his or her aggression onto the stimuli, the response content contains aggressive themes as a result. The examiner then equates these aggressive themes with the personality trait of aggression. When taken to the

extreme, the blank screen metaphor has had two consequences on our approach to projective tests: an overemphasis on response content and an underappreciation for the role of the projective test stimulus and the examination context. By *examination context* we mean the various situational factors as experienced by the respondent. These include the demands on the respondent given the circumstances of the evaluation, the implicit and explicit consequences of the examination, and the interaction between the examiner and respondent.

The blank screen metaphor suggests that the only necessary components to projective test stimuli are ambiguity and a lack of structure. These components are thought to facilitate response content, that is, the free expression of the respondent's internal world. The more ambiguous and unstructured the stimulus, the more it was presumed that the personality would be directly expressed in the response. Historically, this simplistic view has led to an emphasis on response content and to the interpretive viewpoint that the test was equivalent
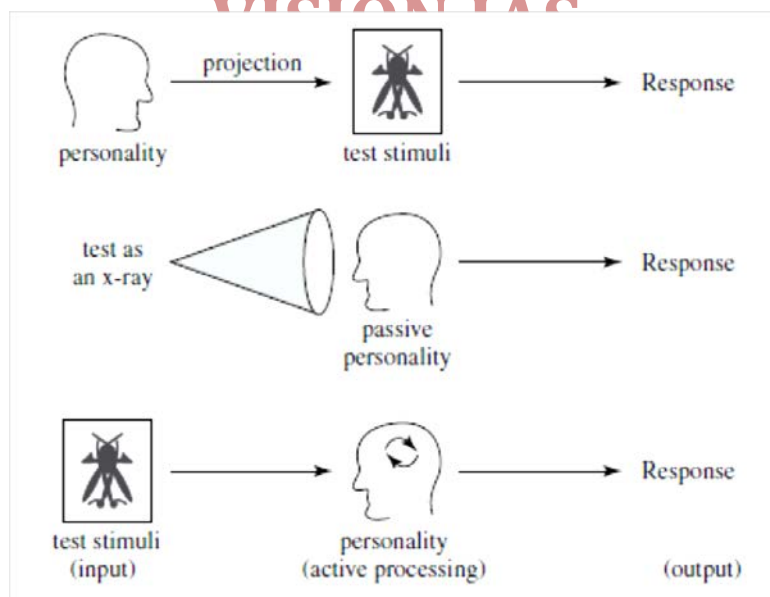
Figure 23.1 Panel A: The theoretical model of the response process as suggested by the blank screen metaphor; Panel B: The theoretical model of the response process as suggested by the X- ay metaphor; Panel C: The proposed problem-solving model of the response process.

to or symbolized an internal response or reality (Murstein, 1963). Aspects of test responses are often seen as symbolic of and equivalent to personality and constituted the basis for grand interpretations. Figure 23.1 presents a schematic for this and other models.

However, increasing the blankness (so to speak) of the screen by increasing the ambiguity of the stimuli does not necessarily produce more useful or valid information. Research into the relationship among amount of ambiguity, structure of pictorial stimuli, and test validity has not led to consistent findings (Murstein, 1961, 1963, 1965). For example, the blank TAT card produces relatively conventional responses that are less revealing of the individual than are the rest of the cards, all of which include a picture of either a person, a group of people, or some other scene. Moreover, eliminating the more recognizable and salient visual aspects of the Rorschach stimuli (what Exner, 1996, called the *critical bits*) does not lead to more productivity. In fact, the available research supports the view that the suggestive aspects of the stimulus, rather than the lack thereof, are what is important. Empirical data clearly demonstrate that the physical stimulus is crucial (Exner, 1974, 1980; Murstein, 1961; Peterson & Schilling, 1983).

What we know about Herman Rorschach's work in developing his test attests to the fact that it is not ambiguity or lack of structure that contributes to the test's usefulness. It appears that each stimulus plate was designed to contain visually recognizable forms, or critical bits, along with some arbitrary components (Exner, 1996, 2000). Rorschach may have included the arbitrary contours to interfere with the processing of these suggestive, recognizable forms. The plates were carefully chosen, drawn, and redrawn so that

many versions existed before Rorschach finalized the designs. Anyone who has ever made inkblots has found that most products look simply like inkblots and are not suggestive of other forms or objects. Thus, it seems that the stimulus plates were intended to be provocative to respondents while also being just unclear enough to engage respondents' problem-solving skills. This inconsistency between the recognizable or suggestive components of the stimulus plates and the more arbitrary forms is critical because it constitutes a problem to be solved. In this sense, projective test stimuli have a clear purpose: to present the respondent with a problem-solving task. For example, a major part of the Rorschach projective task is to reconcile visual and logical inconsistencies among blot details and between the blot and the object (or objects) seen. It is the idiosyncratic ways in which respondents solve the problem, rather than merely the content they project onto a blank screen, that reveals useful and valid information. Thus, understanding projective stimuli as blank screens, rather than as problems to be solved, is a fundamental misconception about projective tests that can lead to inaccurate interpretations of test behaviors.

The X-Ray Metaphor

Another common metaphor is that of an *X-ray machine*. In this metaphor a projective test acts as an X-ray of the mind, so to speak, that allows the interpreter to observe directly the contents of the respondent's mind (see Figure 23.1). Both Frank (1939/1962) and Murray (1938) mentioned this image in their seminal work so that it has historical precedents. However, like the blank screen metaphor, the X-ray metaphor leads to a focus on response content and the way in which the content directly represents personality. More importantly, the X-ray metaphor diminishes the role of the respondent in the response process.

Examining Frank's (1939/1962) original work allows one to achieve a more adequate understanding of his purpose for using the X-ray metaphor. When Frank first used it, he compared learning about personality to the then-current technologies in medical and physical science that allowed one to study internal anatomical structures through non-invasive techniques. However, Frank included a critical distinction between projective tests and medical tools, a distinction that is typically excluded from today's common understanding of the X-ray metaphor. Frank noted that personality, unlike the target of an X-ray machine, is not a passive recipient of attention. In responding to projective test stimuli, personality does not simply cast a shadow of its nature onto a plate. Rather, Frank contended that *personality is an active organizing process*. Despite having been written more than 60 years ago, Frank's ideas reveal a complex and informed perspective on personality, one that is especially relevant to understanding the nature of projective testing:

Personality is approachable as a process or operation of an individual who organizes experience and reacts affectively to situations. This process is dynamic in the sense that the individual personality imposes upon the common public world of events (what we call nature), his meanings and significances, his organization and patterns, and he invests the situations thus structured with an affective meaning to which he responds idiomatically. (1939/1962, p. 34)

Frank went on to describe personality as a "dynamic organizing process." He contrasted this subjective, synthetic, dynamic process of personality to the objective, external, concrete reality of the world, including the host culture's shared conventional experiences. In Frank's view, the world of culture also influences the personality and its understanding of the external world but cannot account for personality processes and behavior.

Later in the same paper, Frank described projective techniques as essentially inducing the activity and processing of the personality:

In similar fashion we may approach the personality and induce the individual to reveal his way of organizing experience by giving him a field (objects, materials, experiences) with relatively little structure and cultural patterning so that the personality can project upon that plastic field his way of seeing life, his meanings, significances, patterns, and especially his feelings. Thus, we elicit a projection of the individual personality's private world because he has to organize the field, interpret the material and react affectively to it. More specifically, a projection method for study of personality involves the presentation of a stimulus situation designed or chosen because it will mean to the subject, not what the experimenter has arbitrarily decided it should mean (as in most psychological experiments using standardized stimuli in order to be

"objective"), but rather whatever it must mean to the personality who gives it, or imposes it, his private, idiosyncratic meaning and organization. (1939/1962, p. 43)

These quotes make it clear that the respondent's organizational style and affect are critical to the projective testing process, and that the process involves more than simply adding content to a stimulus field. Moreover, unlike self-report tests, projective test stimuli give respondents an opportunity to express their organizational styles and affect. Thus, a projective test allows the examiner to observe personality in action with cognitive, affective, interpersonal, and meaning-making activities.

The Need for an Informed Conceptual Framework

This critical review of traditional metaphors and models for projective testing points to their serious shortcomings and oversimplifications. In contrast to a blank screen, projective stimuli are more like problem-solving tasks. In contrast to a passive personality that unknowingly projects itself onto a blank screen or that is examined with X-ray vision, personality in projective testing is seen as a much more active, organizing, and selective process. Perhaps the most accurate portrayal of projection is that the personality does not project light onto the blank screen of the test, but rather, the test projects itself through the active organizing process of the personality to the response. In other words, the individual's personal characteristics are observable in the refracted light—that is, the manner in which the person responds to the test. In sum, there is a need for a broader and more informed conceptual framework for understanding projective testing.

From comparisons between the overt stimuli and response, the interpreter infers the covert personality process. This input-processing-output sequence is the essence of our model for projective testing and is presented in the next section. Such a framework goes beyond projection and response content by embracing a problem-solving perspective.

INTERPRETIVE AND CONCEPTUAL ISSUES

This chapter highlights important characteristics of the projective stimulus situation and the integration of organizational and self-expressive components in the response process. In turn, these factors induce characteristic patterns and methods of interpretation. Projective test responses emphasize synthetic and individualistic approaches to interpretation (Rabin, 1981). In practice, those psychologists who are more inclined to emphasize the complexities of the individual are probably more inclined to use projective tests. The section on free response format outlined the individual or idiographic component of projective test responses. This characteristic of the projective test data induces a similar focus on individual or idiographic approaches to interpretation.

Synthetic, Configurational Interpretation

As established in discussing the projective test stimulus situation and the response process, projective testing accesses multiple dimensions and allows one to elaborate on hypotheses derived earlier in the interpretive process. These factors induce the interpreter to adopt a synthetic or configurational approach in formulating interpretations (Stricker & Gold,1999). Projective test data present connections and associations among various characteristics from different domains. In the TAT, for example, we can associate a cognitive slip or a problem-solving failure with the sexual or intimate themes stimulated by a particular card when such themes are mentioned in (but not meaningfully integrated into) a jumbled and unrealistic story. In terms used earlier in this chapter, projective test results bridge self-expressive, organizational, and response-set domains. One score or response parameter is analyzed in its relationship to another or in relationship to moderator variables and collateral, non-test variables. Temporal, spatial, and language factors—that is, when and how behaviors occur—allow interpreters to identify how various content and organizational aspects of an individual work together, how they interrelate, and how they may interact with different environmental conditions. Advocates of projective testing are not interested in isolated bits of behavior but study how it comes together in a whole person within a life predicament (Stricker & Gold; Viglione, 1999; Viglione & Perry, 1991; Weiner, 1999; Weiner, in press.) Projective test data assist us in putting the person in his or her life context and help us to understand the relationship

78

©VISION IAS
www.visionias.wordpress.com
Contact: +91-9650617807, +91-9968029039

©Numerons
www.numerons.com
Contact: +91-9560344245

between internal issues and the individual's biography and background. Murstein (1963) called this "configurational dominance" (p. 4). This synthetic or configurational approach can be attributed to Gestalt psychology, field theory, and psychodynamic influences on projective testing (Frank, 1939/1962).

Among the connections made possible by projective tests are those between personality factors and cognitive functions. Looking at projective tasks as problems to be solved allows the integration of non-intellectual with intellectual issues. From a configurational point of view, the relationships of abilities to affects, interests, motivations, and discriminative stimuli are addressed. A related advantage of projective tests is that they allow the examiner to make inferences about motivation. Other performance tests (i.e., ability tests) assume and attempt to induce optimal motivation. In real life, however, motivational variation is crucial to understanding personality and behavior.

We can conclude then that part of the utility of projective assessment—or in more concrete terms, its added value relative to self-report tests—is that it provides meaningful connections among different characteristics to enable an understanding of the whole person. Because the individual respondent produces the constructs and their interrelationships with the responses to the projective test in the free-response format, we know that the configurational information is relevant and possibly unique to the individual being assessed. In synthesizing the picture of the individual from a projective-testing perspective, one constructs or builds an integrated picture of the person within a situation. Extrapolating from Kaufman's work on *Intelligent Testing with the WISC-R* (1979), each construction, each person's theory, is different in terms of concepts and relationships among concepts. Reflecting this uniqueness of each individual, projective testing can produce a different theory for each respondent. Along these lines, an important phenomenon is that projective testing often reveals remarkable aspects or concerns that become important hallmarks or organizing features in understanding an individual. Accordingly, assessment report writers often excerpt quotes from sentence-completion responses or responses from other tests to communicate vividly the respondent's experience in the respondent's own words (Holaday, Smith, & Sherry, 2000). Invariably, this synthetic and constructive approach leads to discovering contradictions among test data. Resolving these contradictions often provides added insight into the individual. Recognition of contradictions (e.g., depressed but overly impressed with the self) is based on nomothetic notions. In other words, we see depression and self-importance as being contradictory when we conceive of them as abstract concepts. Within a single individual, these characteristics need not be contradictory. We find that real people appear to possess both overly negative and overly positive views of the self. Positive views may ward off negative views. Positive views may arise in one situation but not another; among children, for example, positive views may arise in academic situations but not at home. It follows then that the inevitable contradictions among projective test data induce, if not necessitate, a dynamic view of individuals. This dynamic view entails opposing forces operating in the behavior, affect, motivation, and cognition in a way that reflects the opposing excitatory and inhibitory organization of the nervous system.

Psychological Testing, Not Psychological Tests

As suggested by the early leaders in assessment (Frank, 1939/1962; Meehl, 1945/1980), the difference between projective and so-called objective tests is not so much in the tests themselves but in the interpretive approach. This difference in approach is induced by the data—by their complexity and richness and their relevance to the individual. Projective tests induce an individualistic, synthetic and configurational, and constructive approach to interpretation that incorporates a view of the individual as embodying contradictions that might be explained dynamically. This approach also involves affects and interpersonal issues. In turn, those holding such a view of interpretation are probably more inclined to use projective tests. The interpreter is involved directly in the assessment administration process, because the individualistic and configurational issues are best known and explored by an active interpreter in a relationship with the respondent. In summary, one's preference for projective tests may largely reflect a philosophical approach to human nature so that it may be more appropriate to talk about projective testing rather than projective tests.

Self-Disclosure and Response Sets

"It would be very unsafe, unwise, and incorrect to assume that a patient either can or wants to present all aspects of his or her personality fully to the examiner" (W. G. Klopfer, 1981, p. 259). Indeed, this is a central problem that clinicians have struggled with in the practice of assessment over the years. Surveys on assessment practice have not explored the extent to which this unsafe assumption is made implicit in the interpretation of self-report personality tests. In the early part of the twentieth century, projective testing grew out of the practical need to access what the individual may be unwilling or unable to communicate directly (Frank, 1939/1962; Murray, 1938; Murstein, 1961). This is the fundamental challenge or paradox in assessment: What is most important to know is often what the person is least willing or able to divulge. To uncover what the respondent may be unwilling or unable to divulge, projective tests go about the task of accessing behavior and problem solving indirectly. Task instructions (e.g., "Tell a story," "Explain what this might be," "Complete the sentence") distract the respondent from the interpretive goals of the examination. Projective tests are attempts to access the private world of the individual, to get to it more efficiently than through other means (Frank, 1939/1962; Viglione & Hilsenroth, in press; Viglione & Perry, 1991).

The reactions to pressures to self-disclose in an indirect stimulus situation are not captured neatly within individual scales on any test. Operating in every individual, idiosyncratically, is a conflict between pressures for self-disclosure versus those for self-protection. This conflict involves (a) a willingness and an ability to be self-revealing versus (b) rational and irrational concerns about negative consequence to self-disclosure, accompanied by (c) a motivation to create a favorable impression. Examination of the nuances of self-revealing behaviors and attitudes to testing in the context of the relationship with the examiner allows us to examine this *struggle over self-disclosure.*

The examiner's strict adherence to his or her own training and to test administration principles, along with careful observation of the respondent and of the respondent's own self-observations, are necessary to manage and observe the respondent's struggle over self-disclosure. For example, in constructing the Rorschach Comprehensive System and in its most recent modification, Exner has gone to great lengths to minimize and to systematize examiner and contextual influences (Exner, 1974, 1993; Exner et al., 2001). Moreover, being sensitive to and evaluating these influences can help one assess their impact on the test findings and inferences (Schafer, 1954). However, the influence of conflicts about self-disclosure and response sets cannot be eliminated. Projective tests offer an opportunity to observe, identify, and characterize these conflicts as a part of the ongoing interaction between the personality and the stimulus situation.

Interpretive Implications of the Pressure to Self-Disclose

The pressure to self-disclose within the projective test stimulus situation leads to a number of interpretive issues. Accordingly, studying and characterizing the response style of the individual is a crucial interpretive goal in all assessment. Response set is an important and complex moderator variable that should be scrutinized in each assessment through observation and analysis of all test data and collateral information. Test findings should be interpreted differently as a function or response set so that the response set acts as a moderator variable for interpretive purposes and validity (Meyer, 1999).

More explicitly, within the interpretive process, results from projective testing can be characterized along the dimension of self-protection versus self-disclosure. Stereotypic, brief test protocols, or poorly or quickly executed productions with insufficient effort (e.g., in drawings) can be seen as attempts to suppress or resist pressure from the examiner to reveal the self. Thus, some test findings may have more to do with how the respondent protects him- or herself or suppresses, defends against, or avoids self-disclosure. Looking at these efforts as a moderator variable, such self-protective test protocols may lead to an underestimate of personality tendencies and weaknesses and to false-negative findings. From a behavioral perspective, this response set can be seen as an attempt to suppress or defend against self-disclosure. In such cases, the test findings do not survey the full array of personality processes and features, so that they may not reveal the personality as a whole. Moreover, these self-protective or suppressive response sets can result in inconsistencies among projective test data, self-report findings, and collateral information (Meyer, 1999).

On the other hand, longer, complex test responses may represent an effort to self-disclose or to express or engage fully in the examination. Such records survey the personality more fully. Alternatively, some overly long and involved test records may represent an effort to present the self in a positive light by demonstrating to the examiner one's talents and problem-solving skills (Viglione, 1996). Nevertheless, too much productivity on any projective test may be associated with overestimation of pathology and false-positive results (Meyer, 1993; Murstein & Mathes, 1996).

It has been well established that response sets vary along this *self-protection/self-disclosure* or *suppressive-expressive* continuum, and that this continuum acts as an important moderator variable in assessment interpretation. Self-report instruments such as the MMPI and the Personality Assessment Inventory (PAI) contain response-set measures such as validity scales and moderator variables. These scales are most useful in measuring the quantitative dimensions of response set. Projective test data are instrumental in individualizing and identifying nuances and complexities in that response set. For example, sentence-completion methods illuminate individual styles, worries, motives, and interests in presenting one's self in an overly positive or negative manner. In that sense, projective testing adds content to what we might learn from the validity scales of an MMPI. Response sets have implications beyond the interpretation of a given test protocol. Attitudes toward self-disclosure/self-protection are fundamental issues in any change process, be it within a clinical, forensic, or organizational setting. Accordingly, the respondent's engagement in testing has implications for motivation to self-disclose in response to interventions in the real world. Similar issues emerge in assessment of the risk of dangerousness. In these contexts, respondents' attitudes toward assessment may also resemble attitudes toward cooperation with management of their risk. Accordingly, these attitudes as a component of the response set are critical assessment targets, and need to be observed closely in assessments. Response set is important, not only as a mediator and discriminative stimulus for test validity, but as a target of assessment in and of itself.

Extreme response sets sometimes emerge as malingering and feigned presentations. For respondents to projective tests, adopting such a response set is quite challenging because of the complexity of the stimulus situation, the active role of the examiner, and the freedom allowed within the test responses. In general, malingering or faking successfully may be more difficult to achieve in projective testing than in self-report testing. A study by Shedler, Mayman, and Manis (1993) reveals that in self-report a substantial portion of respondents may incorporate this false-positive bias in their response styles so as to obscure these tests' sensitivity to problems. These data suggest that projective tests may more accurately describe these individuals' functioning. As for individual tests, research suggests that in some respects Rorschach is more resistant than self-report to response manipulation (Bornstein, Rossner, Hill, & Stepanian 1994; Viglione, 1999).

Nevertheless, the broad claim that the respondent has no control over the content of projective tests is a myth that does not withstand logical and empirical scrutiny. Accumulated research on faking and experimentally induced response sets suggests that a respondent can control content to some extent on many projective tests, including the Rorschach. For example, aggression and sexual content themes, but not dependent and many other themes, are routinely subject to considerable control (Exner, 1993; Viglione, 1999). On the TAT many themes are relatively easily controlled (Holmes, 1974; Murstein, 1961).

Test or Method?

Another long-standing controversy concerns whether projective instruments are actually tests or merely methods or techniques. A *psychological test* can be defined as a standardized administration with an interpretive system that is quantified and subjected to scientific validation. In contrast, a *method* is defined as a systematic way of collecting behavioral observations. Both a test and a method may produce valid interpretations. Within a method, the techniques and strategies of interpretation, rather than quantities produced by scales, would be subject to scientific verification. An example of the use of a projective instrument as a method would be the recognition that completing the sentence stem "I was bothered by" with the written phrase "the way you looked at me when I was putting the blocks together" may have special interpretive significance for the interpretation of Block Design and interpersonal performances. Asking a respondent what he or she had in mind when endorsing "I have two personalities inside of me" would be an example of using a self-report test as a method. Thus, both self-report and projective

instruments could be used as methods. In fact, one might argue that using either of them as a method enhances interpretation.

The Method Argument

These issues have been addressed in the literature. For example, Weiner (1994) published an article on the Rorschach that restimulated considerable controversy about its status as a test versus a method. He suggested that the Rorschach was foremost a method because the instrument is a means of collecting information about how people structure their experiences, express themselves, and interact affectively and interpersonally. It could not be reduced to a single quantification of any specific dimension (i.e., to a test). Similarly, B. Klopfer, Ainsworth, Klopfer, and Holt (1954) advocated for calling the test a *technique,* so that individualistic processing could be emphasized. From a more extreme, but current, viewpoint, Aronow, Reznikoff, and Moreland (1994) focus on response content and regard the Rorschach as a *structured interview*. Most practitioners do not score the TAT, and Little and Schneidman (1955) described it as a "sample of verbal behavior." Earlier, Tomkins (1947) had declared that the TAT was a systematic methodology for personality study—not a test itself. Finally, early memory, sentence, and drawing tasks are routinely used as methods without scoring to collect behavioral observations and personal productions. Advocates and critics use the term "method" for different reasons. Some advocates of projective testing support the term *method* for these projective procedures. Beyond precision of language, they are concerned that essential qualitative and descriptive information will be excluded from consideration if this information is not captured in formal scoring. Critics of projective testing endorse the term *method,* claiming that the non-quantified components are not worthy of consideration. This extremist view excludes from consideration response nuances and connotations, test behaviors, and emotional expressions, as well as the interaction between examiner and the respondent. These characteristics constitute important empirical and objective observations. They are the essence of behavioral assessment and are not captured within the reductionistic view that only test quantities be considered. In cases in which they are relevant (i.e., related to other hypotheses firmly grounded in test-based inferences), these behavioral and empirical data derived from using projective instruments as methods must be included in the interpretive process.

Methods, Clinical Utility, and the N-*of-1 Problem in Assessment*

How does one fit a group or statistical concept, such as depression or aggressive risk, to an individual and describe its idiosyncratic experience and function within that individual? From a statistical viewpoint, if it is highly likely that a person is depressed based on a score, how do we confirm the presence of depression in the particular individual we are evaluating? These questions reflect the *N-of-1* problem in assessment—that is, the challenge of applying abstract, group-derived constructs and measurements to a single individual. Within individuals, constructs such as aggression or depression exist only in idiosyncratic forms. Accordingly, within a projective test protocol, idiosyncratic evidence of depression may serve to confirm and individualize a person's expression of depression. In this way, using projective instruments as a method helps address the *N*-of-1 problem in assessment by contextualizing and individualizing abstract concepts and group data.

This *N*-of-1 problem is often framed in terms of the distinction between nomothetic and idiographic science and knowledge (Murstein, 1963). *Nomothetic science* addresses laws and abstractions applicable across circumstances and individuals. Within psychology, it would be associated with group psychological principles, constructs, and data across individuals and situations. These abstractions may not actually exist in any individual case but are hypothetical constructs created for the purpose of explaining and summarizing relationships among groups of individuals. In contrast, *idiographic science* is concerned with understanding a particular event, such what led to a particular historical event or decision—in other words, how and why something happened. The aim of assessment, to characterize a unique individual within a life context, is an idiographic goal. Certainly, nomothetic science, methods, and comparisons are critical and necessary to address this goal, but not sufficient to achieve it fully. Idiographic and configurational information from a method perspective is necessary to address the uniqueness of each case. Thus, projective test data and observations are helpful in translating group, nomothetic, or actuarial data to the individual *N*-of-1 case.

In terms of clinical utility, using an instrument as a method offers considerable advantages over using instruments strictly as tests. Observations and inquiries can be adapted to address any purpose. One cannot imagine all of the questions that will come up in an assessment. Thus, one method might replace many tests, offering considerable efficiency and cost savings. The superior status of tests in terms of the validity of a specific interpretation relies on stringent research validation of the test for that particular purpose. On the other hand, it is impossible to develop, research, and master a test for every purpose. Accordingly, projective methods, interviews, and observations are always necessary for a comprehensive assessment, lest we give up all idiographic assessment goals.

At the broadest level, research supporting the validity of a method addresses whether a projective procedure can produce valid and useful information when projective instruments are used in the standard ways. The research clearly supports the conclusion that the major projective instruments (inkblot perception and representation, storytelling, sentence completion, early recollection, and figure drawing) can yield valid and useful information. On the other hand, the limits of these methods and the limits of data they produce are not fully appreciated by some projective-test advocates. Further research needs to identify the types of inferences and generalizations that can be made about particular personality processes and from which types of data.

Conclusion and Recommendations for Tests and Methods

Projective instruments, like all psychological tests, can function both as methods and as tests. In both roles, they should be administered in a standardized fashion. When they are used as tests, one relies on quantification, measurement against comparison-group data, and pre-established criterion validity. These factors lead to a strong scientific foundation for the interpretation of tests. Because of the less sturdy support for inferences based only on using the instruments as methods, inferences derived from methods need additional support from other sources. Within a given assessment, this support can be accomplished in terms of addressing hypotheses that have derived initial support form data encountered earlier in the assessment process. For example, in established cases of depression, the TAT may yield important information about the idiographic experience of that depression and its interpersonal correlates. Early memories may provide subjective and experiential patterning associated with this depression. If we establish from a self-report test that the respondent is describing him- or herself in an overly positive and defensive fashion, an examination of sentence-completion results and observations about the examiner-respondent interaction may lead to important information about the character and motivation associated with that defensiveness. If new hypotheses emerge from method data, they must be supported by other data from other observations and findings, in a way that we would not require for an interpretation from an instrument used as a test. Thus, when these procedures are used as methods and not tests, they should generally be used as ancillary, or elaborative, procedures.

Rorschach Comprehensive System interpretation is a good example of using an instrument as both a test and a method. One first interprets structural quantitative data, then modifies these foundational interpretations with the Rorschach used as a method (i.e., through response verbalizations and behavioral observations). In this way, method findings are used to refine, elaborate, and qualify previously formulated general hypotheses.

Contribution to Assessment Relative to Self-Report Tests

One way to address the question of what projective testing contributes to assessment is to identify situations in which self-report tests do not yield clear and definitive findings. This approach is consistent with the current concerns about incremental validity. Many have noted that projective tests contribute most in contexts in which the person may be unwilling or unable to provide the sought-after information through more direct means (Bagby, Nicholson, Buis, & Radovanovic, 1999; Bathurst, Gottfried, & Gottfried, 1997; Bornstein, 1999; Viglione, 1999). Some might contend that, to some degree, no respondent is able to fully reveal critical information about the self in an efficient manner.

The traditional view, as first elaborated by Frank (1939/1962), is that projective testing goes beyond socially conventional meanings and roles. From this perspective, self-report items, typically a sentence or

phrase, presume a conventional, widely shared understanding of their meaning. In these conventional contexts, individual behavior is best explained by situational phenomena as interpreted with shared cultural norms. Frank contrasted these conventional contexts to situations in which behavior is explained by the individual's unique ways of (a) ascribing meaning to the world and (b) organizing the world. In fact, one's unique ways of ascribing meaning to and organizing the world is the fundamental component of personality, according to Frank. Moreover, they correspond to the self-expressive and organizational components of projective tests addressed earlier in this chapter. Projective tests are designed to access these individualistic functions and thus reveal personality activity directly. This linking of self-report tests to conventional contexts and projective tests to individualistic ones has led some to speculate about the relative contributions of these tests. For example, Hutt (1945) speculated that self-report tests may be valid only when the respondent is willing and able to self-rate on a known dimension. Meehl (1945/1980) disagreed by objecting that although respondents may understand self-report test items differently, such differences are not relevant to validity. He claimed that the validity of a self-report test is not a function of a conventionally, socially prescribed understanding of the test items. Rather, it is a function of empirical relationships with external criteria. This *empirical keying approach* assumes that the content of the item really does not matter, only its empirical relationship with meaningful criteria.

Despite Meehl's (1945/1980) assertions, evidence suggests that what the item means to the respondent does make a difference in the validity of a self-report personality test. On the MMPI, it is well established that obvious items are more valid then subtle items (Graham, 2000; Greene, 2000; Gynther & Burkhart, 1983). In other words, when an item's content is semantically related to the scale on which it resides or the construct it measures, it works better. Also, the largely rationally derived content scales on the MMPI-2 rival the empirically keyed clinical scales in terms of validity (Graham; Greene), again suggesting that item content matters. The current test-development practice is to pick the item pool for content validity (i.e., what the item means to the respondent; American Educational Research Association et al., 1999; Anastasi & Urbina 1996; Morey, 1996). Again, the validity of these scales is partly based on an unequivocal meaning of the item to the respondent. As Frank (1939/1962) asserted theoretically and McClelland et al. (1989) and Bornstein et al. (1994) demonstrated with data, self-report personality tests reveal information about relevant but conventional, culturally prescribed dimensions.

The interpretive implication of all of these data is that self-report personality tests tell us the most about social role–related behavior, how one behaves in the role of a father or in the role of a rebellious adolescent in our society. These tests work best when the examinee translates individual items in conventional ways and when the examinee's response set reflects the host culture's norms. Psychometrically, this occurs when validity scales (e.g., *L*, *F*, and *K* with the MMPI) are near average values. Atypical, unconventional response sets, in terms of excessive defensiveness or exaggeration, reflect unconventional approaches to the tests; and atypical translation of test items, in turn, limits the validity of self-report personality tests (Meyer, 1999). Conversely, projective tests have the most to offer in understanding and predicting behavior outside prescribed social roles and demands across situation and time, as well as for issues that are idiographic, idiosyncratic, or implicit (see Bornstein, 1999; Shedler et al., 1993; Viglione, 1999). These would include environmental contexts or patterns of behavior that are structured by individual personality rather than by social roles and conventions.

PROJECTIVE TEST CONTROVERSY FROM A HISTORICAL PERSPECTIVE

This chapter attempts to clarify many misunderstandings about projective testing. These misunderstandings can also be seen in a historical perspective. Undeniably, historical developments have influenced our understanding of focal psychological constructs, even when we believe that these constructs are grounded in empirical science. For example, as a result of the Wechsler and Stanford-Binet scales, our implicit and conventional understanding of intelligence emphasizes the quantitative perspective at the expense of the conceptual and developmental aspects as articulated within the Piagetian approach. Self-report personality assessment has led us to simplify adult personality into an aggregate of traits demonstrated by subgroups of individuals. Response set or response manipulation has been reduced to quantitative notions about exaggeration and defensiveness (e.g., as defined through the *L*, *F*, and *K* scales on the MMPI). Thus, history and our experience have shaped our views, constructs, and what we consider to be science.

Emerging Clinical Needs versus Scientific Aspirations

Our current views of assessment and the relative values of projective and self-report tests of personality are shaped not only by metaphors and models, but by historical traditions as well. Misunderstandings about projective testing have shaped the ongoing and lengthy controversy about projective tests. It is surprising to learn that the current controversy about the utility of projective tests surrounding the use of these tests has existed since their introduction (Hirt, 1962; Murstein, 1965; Rabin, 1981). The popular academic-scientific position dating back to the 1920s is that projective tests are flawed. Periodically, this view has been a rallying cry of academic psychologists. In the 1920s and 1930s, American academic psychology focused on distinguishing psychology by making it a science with mathematical foundations much like those of physics. It is not surprising that it produced few concepts, facts, and methods applicable to clinical work. At that time, applied work in clinical psychology was largely diagnostic and descriptive in support of psychiatrists' work with individuals with mental disorders. These clinical and practical demands opposed the academic interests in developing the discipline and science of psychology.

The need for personnel selection in the military and evaluation and treatment of consequences of the two world wars further stimulated the practical needs of applied psychologists. More generally, clinicians thought that the individual was lost in the techniques of the so-called mental testers. They wished to recognize the interaction between individual characteristics and "the total life situation which could lead to an adequate description of the person as a functioning human being" (Murstein, 1963, p. 4).

As it has been in the past, *projective testing* continues to be a rallying symbol for those wishing to move beyond the response manipulation in self-report tests so as to understand the individual. Thus, clinical and applied interest and questions that outstrip scientific and academic developments in the field have marked the whole history of assessment. As society changes, this pressure to address advanced and complex questions in everyday practice will certainly persist. Nonclinical psychologists who criticize projective tests may not fully understand the demand society justifiably places on our clinicians and the interpretive usefulness and validity of the behaviors collected and observed by using projective tests.

The Polarized and Moralistic Debate Continues

The controversy about the value of projective persists to this day. The result is that too much of the attention given to projective tests in the literature are polemical and editorial rather than scientific (APA, 1999; Dumont & Smith, 1996; Garb, 1998, 1999; Garb, Florio, & Grove, 1998, 1999; Garb, Wood, Nezworski, Grove, & Stejskal, 2001; Grove & Barden, 1999; Joiner & Schmidt, 1997; Wood & Lilienfeld, 1999; Wood, Lilienfeld, Garb, & Nezworski, 2000; Wood, Nezworski, & Stejskal, 1996, 1997). The end result is that projective and self-report tests are pitted against one another as adversaries. The most recent manifestation of this rivalry is the current application of incremental validity with the unqualified and simplistic assumption that projective tests should increment above self-report tests in regression equations (Hunsley & Bailey, 1999). This position grossly oversimplifies the clinical endeavor (Viglione & Hilsenroth, in press) while ignoring research demonstrating incremental validity for projective tests (e.g., for the Rorschach; Archer & Gordon, 1988; Archer & Krishnamurthy, 1997; Blais, Hilsenroth, Castlebury, Fowler, & Baity, 2001; Bornstein, Bowers, & Robinson, 1997; Cooper, Perry, & O'Connell, 1991; Holzman et al., 1974; Meyer, 2000a; O'Connell, Cooper, Perry,&Hoke, 1989; Perry & Braff, 1994; Perry & Viglione, 1991; Russ, 1980, 1981; Shapiro, Leifer, Martone, & Kassem, 1990; Skelton, Boik, & Madero, 1995; Viglione, 1999; Weiner, in press).

The very name *projective* is subject to these politics and polemics. The persistence of the nomenclature of *objective* for self-report tests, in juxtaposition with *projective,* further distorts the data and viewpoints of psychologists. This dichotomy implies that the virtues of objectivity and psychometric discipline are reserved for the self-report tests. Moreover, *projective* is associated with *subjective* and the cardinal sin of bias when it is juxtaposed against *objective*. Research demonstrations of the incremental validity of projective tests support the view that extra validity is accessed through involving the examiner in the interpretive processes with projective tests. This added validity is obtained even if there is more examiner variability or examiner unreliability with these procedures. The most vexing problem in this debate is that

the views are drenched in ethical and moralistic language, so that the polarized positions are experienced as moral imperatives. This moralism from the academic side has to do with claims of the righteousness of science and empirical foundations. On the other hand, there is some validity to the claim that clinicians using projective tests have historically shown a tendency to overpathologize (Murstein & Mathes, 1996; Viglione & Hilsenroth, in press).

The advocates of projective testing are not immune to similar criticisms. Murstein (1963) correctly pointed out that the seminal articles (e.g., Frank, 1939/1962; Rosenzweig, 1951) have a moralistic tone, with the hero being the idiographic clinician using projective testing to describe and understand the individual in all his or her complexity. The idea that clinical interpretation is an art was described by Levy (1963) as "a romanticism with which some are afflicted that seems to render its victims either insensitive or antagonistic to any attempt at rational analysis . . . [T]hey (proponents of projective testing) rely on 'moralism.' To criticize the 'mental testers,' they use epithets such as 'atomistic,' 'mechanistic,' and 'superficial' " (p. 3).

A. Kaplan (1964) has given a slightly different slant to these polemics in his description of seductive and reductive fallacies. Projective test advocates seduce themselves and others into believing there is always something else, subtle or otherwise, that can be gleaned from the data and applied to useful purposes. Kaplan refers to this belief as the *seductive fallacy*. On the other hand, projective tests critics embrace a *reductive fallacy,* which incorporates the view that science requires that test data incorporate a very limited number of key elements. Furthermore, this belief requires that automated techniques must be involved because clinicians cannot reliably identify these key elements.

As an example of this bias against projective tests, Masling (1997) questioned whether the data supporting the Rorschach would change the minds of the critics, given the persistent history of bias against projective tests. He attributes some of this rigidity, politicization, and bias to the fact that former students, emboldened by their teachers, have become critics of projective tests. Unfortunately, some students of psychology are socialized to believe in simplistic models, such as the blank screen, or in the supposedly unscientific foundation of projective techniques in order to continue the conflict (Viglione & Hilsenroth, in press).Weiner (1996) also observed that the critics had ignored 20 years of empirical support for the test. One can only conclude that these speculations were correct, as the recent debate about the Rorschach has demonstrated. Data and experience suggest that these critics continue to ignore the research supporting projective testing in general and the Rorschach in particular (e.g., Meyer, 1997a, 1997b, 2000a, 2000b; Meyer & Archer, in press; Riethmiller & Handler, 1997; Shedler et al., 1993; Stricker & Healy, 1990; Viglione, 1999; Viglione & Hilsenroth;Weiner, in press).

Recognition of Differences and a Resolution?

The rivalry and controversy about projective and objective personality tests may merely be a manifestation of the conflicts and misunderstandings between clinical and academic psychologists or between practice and science in psychology. A great deal of psychology's time, energy, and intellectual effort have been wasted within this war. Sadly, American psychology has not been able to resolve this dilemma, most likely because it is a basic philosophical and moral disagreement rather than a scientific one. American psychology perseveres under the goal of integration of science and practice, yet those who embrace this vision often hold very different perceptions of this integration. This science-versus-practice debate continues with little understanding or appreciation of the other's point of view, and with little hope for reconciliation and advancement of assessment psychology.

Our goal should be to diffuse this conflict and integrate the strengths of projective and self-report approaches to assessment. As Levy (1963) points out when he calls for systematic empirical and rational evaluations of clinical interpretation, "to this writer's way of thinking, rationality and human dignity are not antithetical to each other; nor are science and art for that matter" (p. 3). Each type of test has strengths and weakness. We are lucky that cognitive, projective, and self-report tests complement each other so well.

Meehl (1945/1980), an advocate of self-report personality testing and scientific psychology, asserted that "there is assuredly no reason for us to place self-report and unstructured types of instruments in battle

order against one another, although it is admitted that when time is limited they come inevitably into a very real clinical 'competition' for use" (p. 302). He maintained that all personality tests can be see on a continuum and can be interpreted with the same principles. As described earlier in this chapter, he placed the difference between projective and self-report within the interpreters and their own philosophies rather than within the instruments. Historically, advocates of projective and self-report personality testing have presented the arguments of the opposing side as being seriously flawed. This destructive rivalry could be replaced with reconciliation and is opposed by the argument that all tests incorporate so-called projective and objective features, as well as empirical characteristics. The basic premises, goals, and activities are different but potentially complementary (Masling, 1997; Meehl, 1945/1980). As Masling concludes his paper on projective testing, Psychology is fortunate to have available two such different means of assessing human behavior, each placing different emphasis on the importance of motive, rational thinking, fantasy, self-reflection, and defense. A wise discipline would value and embrace such differences rather than find fault and lack of respectability in either one. (p. 266)

Adopting some of the perspectives described in this chapter may assist in integrating projective and other approaches for more effective assessment. Most important among them may be adopting the response-process approach, including both problem-solving and behavioral components. One challenge in writing this chapter has been to encompass the great diversity of projective tests under one umbrella. The extant research data and the response-process model itself would suggest that the next step would be to adapt the model to individual tests. This would include developing paradigms to address topographical and experiential similarity, functional equivalence, and personality as problem-solving in real life. The challenge in this research is to access the idiographic characteristics of the individual as validity criteria. This is not a simple manner and may require incorporating qualitative research with more traditional quantitative work. Research should also tackle the international and cross-cultural challenges, since projective testing has great potential in these applications. Every effort should be made to standardize administrations and coding of responses, with the Rorschach Comprehensive System as the model. It is unclear whether we can progress much further by lumping these tests together. Rather, the response process and generalization characteristics for each test can be researched and developed separately. Research in projective testing should address the interpretive process itself. Much more sophisticated clinical judgment studies are needed to make them relevant to clinical practice (Karon, 2000; Levine, 1981; Viglione & Hilsenroth, in press). Such research should include investigations of these instruments as methods.

(b) Discuss A-B-C components of attitude. Describe the procedure of development of a tool for measurement of attitude.

TOPIC: *Attitudes, Values and Interests*
SUBTOPIC: *Components of Attitudes*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

While defining *attitudes,* social psychologists focus on the tendency to like or dislike an attitude object or behavior. That is, attitudes are defined as tendencies to *evaluate* objects favorably or unfavorably (Bem, 1970; Eagly & Chaiken, 1993; Fazio, 1990; Olson & Zanna, 1993; Petty, Wegener, & Fabrigar, 1997; Wood, 2000). Attitudes can be directed toward any identifiable object in our environment, including groups of people (e.g., ethnic groups), controversial issues (e.g., legalized abortion), and concrete objects (e.g., pizza). In fact, the potentially unlimited range of attitude objects sometimes causes confusion about the relations between attitudes and other social psychological constructs. For example, there is conceptual overlap between attitudes and values, which are abstract ideals that people consider to be important guiding principles in their lives (e.g., freedom; Rokeach, 1973; Schwartz, 1992). The *importance* component of values makes them distinct from attitudes (Feather, 1995; Maio & Olson, 1998), because positive attitudes do not imply that the targets are important guiding principles in life. One fundamental attribute of attitudes is that they are subjective—that is, they reflect how a person sees an object and not necessarily how the object actually exists. Consequently, attitudes should be considered a part of the subjective self, which is the stream of thoughts, feelings, and actions that govern how someone lives (James, 1890).

A-B-C Components of Attitudes

Structure of Attitudes

Attitudes structure can be described in terms of three components, affective, behavioral and cognitive. This model is known as the ABC model of attitudes. The three components are usually linked. However, there is evidence that the cognitive and affective components of behavior do not always match with behavior. This is shown in a study by LaPiere (1934).

Affective component: This involves a person's feelings / emotions about the attitude object. For example: "I am scared of spiders". Individuals' evaluations of targets can also be based on how the target makes them feel—that is, on the emotions or affect aroused by the target. Indeed, as noted in this chapter's section on attitude structure, affect sometimes predicts attitudes better than does cognition (e.g., see Esses et al., 1993). Of course, affect and cognition are often (or even usually) consistent with one another because these processes are mutually interdependent (e.g., knowledge can influence feelings, and feelings can guide thoughts). Although affect toward objects can spring from beliefs about those objects, there are a number of processes that can result in affect's becoming associated with an object independently of cognition (i.e., independently of information about the characteristics of the object). These processes are discussed in detail in the section entitled "Low-Effort Attitude Change Processes" in the chapter on attitude change (see the chapter by Petty in this volume), so we only mention them here briefly. One process is classical conditioning, which occurs when a stimulus comes to evoke a response that it did not previously evoke, simply by being paired with another stimulus that already evokes that response. For example, the receptionist at a dental office might come to evoke negative affect for patients who are very fearful of dental work. Although a conditioning perspective on attitudes has been around for many years in social psychology (e.g., Staats & Staats, 1958), the past decade has continued to see very sophisticated studies documenting conditioning effects on attitudes (e.g., Cacioppo, Marshall-Goodell, Tassinary, & Petty, 1992).

A second process through which affect can become linked to objects without necessary cognitive mediation is mere exposure. The *mere exposure effect* (Zajonc, 1968) occurs when repeated, simple exposure to an object (i.e., exposure without reinforcement feedback) leads to more favorable feelings toward the object. For example, an abstract painting that initially evokes confusion might come to be liked over time—simply because the painting is more familiar. The results of several fascinating studies have shown that conscious recognition that stimuli are familiar is not necessary for the mere exposure effect to occur (e.g., Moreland & Beach, 1992), nor, in fact, is conscious *perception* of the object—subliminal exposures can increase liking for a stimulus (e.g., Bornstein & D'Agostino, 1992).

Behavioral (or conative) component: the way the attitude we have influences how we act or behave. For example: "I will avoid spiders and scream if I see one". A third potential source of attitudes is behavioral information—specifically, knowledge of one's previous actions toward a target. This knowledge can influence attitudes through a variety of processes, including dissonance arousal and self-perception processes. From the perspective of dissonance theory (Festinger, 1957), knowing that one has acted favorably or unfavorably toward a target will *motivate* an individual to evaluate the target in a manner consistent with those actions (e.g., Cooper & Fazio, 1984). From the perspective of self-perception theory (Bem, 1972), individuals might logically *infer* that their attitudes are consistent with their actions (e.g., Olson, 1992). Thus, an effect of past behavior on attitudes may reflect both cognitive and affective processes. In a recent paper, Albarracin and Wyer (2000) reported several studies in which they cleverly tested the effects of knowledge about past behavior by leading participants to believe that they had expressed either support for or opposition to a particular position without being aware of it. Because participants had not actually engaged in such behavior, the research tested directly the effects of *believing* that one has behaved in a certain fashion. Results showed that participants reported attitudes that were consistent with the alleged past behavior and that subsequent behavior toward the target also tended to be consistent with the alleged prior action. Thus, behavioral information had a direct effect on attitudes and subsequent behavior.

Cognitive component: this involves a person's belief/knowledge about an attitude object. For example: "I believe spiders are dangerous". One crucial source of attitudes is cognitive information about the target— that is, beliefs about the attributes of the target. Indeed, as discussed in the section of this chapter on attitude structure, beliefs play a prominent role in both major models of attitude content. Knowledge about an object can come either from direct experience with the object or from indirect sources such as parents, peers, and the media. As already noted, attitudes based on direct experience tend to be stronger than are attitudes derived from indirect information. The best-known theory of attitude formation based on cognitive beliefs is the *theory of reasoned action* (Fishbein & Ajzen, 1975), which is an expectancy-value model in which salient (i.e., highly accessible) beliefs are hypothesized to combine additively to form the overall evaluation of the target (attitude toward the target). As noted earlier in the chapter, many researchers have documented a strong relation between attitudes and expectancy-value products (e.g., Budd, 1986; van der Pligt & de Vries, 1998). This model of attitudes is based on a conception of humans as rational, deliberate thinkers who base their attitudes and behavior on information about the positive and negative consequences of various actions.

(Describe the procedure of development of a tool for measurement of attitude.)

Attitude Measurement and Questionnaire Design

A great deal of research has examined methods of designing questionnaires to measure attitudes and other subjective psychological constructs effectively *(see also* ATTITUDE THEORY AND RESEARCH, SURVEY METHODS), and this research provides a basis for a number of recommendations. We shall begin below by reviewing some of the original attitude measurement techniques, which involved elaborate procedures and multiple items. Then we will review the more recent literature on single-item measurement and the many decisions one must make in designing such measures.

Early Attitude Measurement Methods: Multiple Item Indices

A number of elaborate attitude measurement techniques were developed beginning in the 1920s (for reviews, *see* Dawes & Smith, 1985; Mueller, 1986; Summers, 1970). One such technique was developed by L. L. Thurstone. In his classic paper, "Attitudes can be measured," Thurstone (1928) argued that attitudes toward objects could be gauged in ways similar to those used to assess perceptions of sensory stimuli such as light and sound. Although this idea does not seem to be particularly revolutionary today, it marked a bold departure from the dominant behaviorist tradition of the time, which held that latent psychological constructs were not legitimate topics of scientific inquiry or assessment.

Thurstone developed several techniques for measuring attitudes, the most popular of which was the method of equally appearing intervals (EAI). EAI scales are developed by generating a large pool of statements, each of which expresses some degree of positivity or negativity toward a target object. A group of judges then classifies each statement into one of eleven categories according to how much positivity or negativity the statement reflects toward the object. The categories are then numbered from 1 to 11, representing increasing positivity, and a scale value is assigned to each statement by computing the median or mean rating of the statement (on the 1 to 11 scale) across the judges. A final item set is then constructed by selecting one statement to best represent each of the 11 intervals. This final set of items can then be administered to respondents, instructing them to indicate with which statements they agree. The median or mean scale value of the statements with which each respondent agrees serves as the index of the respondent's attitude.

Although the EAI method is highly reliable, the extensive preparation necessary to construct EAI scales spurred researchers to investigate other, simpler methods. One popular one proposed by Rensis Likert is the method of summated ratings. This method begins by generating a large number of statements reflecting positivity or negativity toward the target object. Respondents then indicate the extent of their agreement or disagreement with the statements using 5-point agree/ disagree scales. Next, responses to these items are summed to create an overall score for each respondent, and the correlation of each item with the total score is computed. Items that correlate poorly with the total score are deleted, and the remaining items are used

to yield a final index of the attitude. Thus, the method of summated ratings does not require a separate group of judges to rate items prior to administering the final scale to the sample of interest.

Another measurement method traditionally used to assess attitudes towards ethnic groups is the method of social distance. According to this method, respondents are given a series of statements reflecting increasingly proximal degrees of social contact with members of a target social group (e.g., shopping in a store where members of the target group shop versus having a member of the target group over for dinner). The respondent's attitude is assessed by determining the most proximal behavior the respondent is comfortable with.

Another method of measuring attitudes is the semantic differential, developed by Charles Osgood and his colleagues. These investigators argued that attitudes can be measured by asking respondents to rate an object on bipolar scales anchored by pairs of adjectives reflecting an evaluative, positive-negative dimension (e.g., good-bad, wise-foolish). Responses to these scales can then be summed to arrive at an overall attitude score for each respondent. Because the same adjective pairs can be used for nearly any attitude object with little if any pretesting, this method has been quite popular among social scientists.

Contemporary Attitude Measurement: Designing Single Item Measures

Unfortunately, it is often not practical to measure attitudes using one of the above methods, because they involve multiple items per attitude and sometimes involve elaborate pretesting. Especially in surveys in which attitudes towards dozens of objects are assessed, it is not feasible to obtain multiple measures of each attitude .. Consequently, researchers have turned increasingly often to single item measurement approaches that they believe best reflect the underlying attitude.

An advantage of the multiple item approach is that the particular characteristics of any particular item are unlikely to have a substantial impact on the results of an assessment procedure. But when one relies on only a single item, its characteristics can potentially have tremendous impact on one's conclusions. Therefore, one must design single item measures very carefully in light of one's research goals. Below, we will review the existing literature on designing single items and summarize some of the recommendations supported by these studies (for more details, *see* Himmelfarb, 1993; Krosnick & Fabrigar, in press; Schuman & Presser, 1981).

Open Versus Closed Questions

One of the first issues that a researcher must confront when constructing an item is whether to use an open-ended or a closed ended question. Closed questions provide a list of response options among which a respondent must choose, whereas open questions allow respondents to answer in their own words. Although researchers have typically used closed questions due to the ease of administering and coding them, recent research has suggested that this may sometimes be at a cost to data quality.

One drawback of closed questions is that respondents usually limit their responses to the offered alternatives, even if their optimal answers are ones not mentioned in the question. For example, if asked "What is the most important problem facing the country today: unemployment, inflation, the government's budget deficit, or some other problem?", nearly all respondents will choose one of the first three options rather than generating alternatives. Yet when these same individuals are asked an open-ended question, they are likely to generate a much larger set of important problems. It is possible to avoid this problem, however, if one builds lists of closed question response alternatives based upon pretesting with open-ended questions.

One concern about open-ended questions is that they may disproportionately reflect concerns or feelings that happen to be on the minds of respondents at the time a question is asked, rather than tapping deeper, more considered beliefs or attitudes. According to a number of studies, momentary SALIENCE of considerations does indeed influence responses to open questions, but salience does so to about the same extent with closed questions. Therefore, this concern does not seem to be a basis for shying away from open questions.

Finally, some research in this area has investigated whether closed questions are more likely than open questions to elicit vacuous responses from respondents who actually have no opinion toward an object. This work has indeed found that respondents are more likely to give answers to closed questions than to open questions when the target is a fictitious object, toward which respondents could not have an opinion. Presumably, the ease of responding to closed questions encourages respondents to provide answers even when these answers are not meaningful. Taken together, then, this literature suggests a number of advantages to open-ended questions, so they may be particularly useful for some attitude measurement tasks.

Rating Versus Ranking

When one uses a closed question format, one must choose between a rating format and a ranking format. Rating formats require respondents to report the absolute magnitude of a psychological construct along a continuum (e.g., ranging from "like a great deal" to "dislike a great deal"). Rankings require respondents to order a set of objects according to some criterion (e.g., most preferred to least preferred).

Rating formats are more common in attitude research, presumably because rankings have a number of inherent disadvantages. For example, they yield ordinal and ipsative data, which are less informative and harder to analyze than the interval level data provided by ratings. Furthermore, rankings are a great deal more time consuming and difficult for respondents to complete. Nonetheless, there is some evidence that rankings yield more valid data. Specifically, rankings yield more reliable data than ratings, produce more interpretable factor solutions, and have greater criterion validity. Furthermore, ratings appear to be more susceptible than rankings to response set biases, due to a failure of some respondents to make fine distinctions among objects in rating tasks. Therefore, rankings have a number of psychometric advantages when one is interested in comparisons of attitudes across sets of objects.

Number of Scale Points

When using rating scales, one must decide how many scale points to use for each item. Popular single-item attitude measures have ranged from as small as 2-point yes/no or agree/disagree scales to ones as large as 101 points. On one hand, using scales with more points may allow for greater precision. On the other hand, there may be limits to respondents' ability to make fine discriminations, so increasing the number of scale points beyond a certain point may enhance random error rather than enhancing information acquisition. In fact, the various empirical studies on this matter indicate that scales with 5 to 7 points seem to be both more reliable and valid than scales with more or fewer points (*see* Krosnick & Fabrigar, in press).

A related issue is the question of whether one should use rating scales with odd numbers of points (i.e., including a clear midpoint) or whether to use an even number of scale points. Including a midpoint allows respondents with neutral opinions to report them rather than arbitrarily indicating either a positive or a negative attitude. It is also possible, however, that offering the midpoint may reduce validity by providing a response alternative that is easy to select without much thought and therefore discourages respondents from expending the cognitive effort to report a more substantive view based upon their beliefs about an object.

Empirical research on these matters currently supports four conclusions (*see* Bishop, 1987; Krosnick & Fabrigar, in press; Schuman & Presser, 1981). First, respondents are unlikely to report neutral opinions if a midpoint is not explicitly provided, and they are much more likely to report such opinions when a midpoint is offered. Second, the distribution of positive and negative attitudes expressed can sometimes be very different depending upon whether a response scale includes a midpoint or not. Therefore, the decision regarding whether or not to offer a midpoint may well have a significant effect on substantive research conclusions. Third, a couple of initial studies suggest that the validity of attitude reports is enhanced when a middle alternative is provided. However, additional evidence suggests that middle alternative selection may in fact reflect a desire on the part of some respondents to avoid the cognitive work necessary to formulate and report substantive opinions. Consequently, it is difficult to say at this point whether including a midpoint is desirable.

Branching

Nearly all single-item attitude measures ask respondents to place themselves on a scale ranging from favorable to unfavorable, thus reporting attitude direction and extremity in one step. However, the difficulty of administering long scales during survey interviews over the telephone has led some researchers to employ a branching approach. Accordingly, respondents are first asked whether their attitude is positive or negative or neutral (i.e., direction only). Then, in a follow-up question, respondents expressing positive or negative attitudes are asked how extreme those attitudes are. Respondents who express a neutral attitude initially can be asked a follow up about whether they would lean in a positive or negative direction. This is called branching because the wording of the follow up question varies depending upon which initial answer a respondent provides. A number of recent studies indicate that this decomposing of the reporting process into 2 steps enhances the speed and ease with which respondents can report their attitudes as well as the reliability and predictive validity of those reports (Krosnick & Berent, 1993).

Verbal Versus Numeric Labels

It is quite common for researchers to design long rating scales with verbal labels only on the endpoints (e.g., "like a great deal" versus "dislike a great deal") and· perhaps also at the midpoint (e.g., "neither like nor dislike"). Thus, the precise meaning of the other scale points is left implicit. However, a great many scaling studies have been done in recent years to quantify the meaning that people attach to such modifiers as "a great deal," "somewhat," "a little," and so on, so it is conceivable that consensually interpreted and evenly spaced verbal labels can be attached to all points on ratings scales (except, of course, very long ones). A number of studies suggest that this reduces ambiguity in the meanings of those scale points and thereby enhances reliability *(see* Krosnick & Berent, 1993).

No-Opinion Filters

Regardless of whether open or closed questions are used, a researcher must decide whether to include a no-opinion filter. Such filters sometimes involve asking respondents if they have an opinion toward an object and then only asking what that opinion is if they say they do in fact have one. Alternatively, filtering sometimes involves simply including a "don't know" or "no opinion" category as one of the response options in a single attitude question. Not surprisingly, many more respondents indicate that they have no opinion when a filter is included than when no filter is offered. Again, however, it is not clear whether respondents select a "no opinion" response because they truly have no opinion or because they simply want to avoid the cognitive work necessary to formulate and report an opinion. Because including a no-opinion filter can dramatically alter the proportions of favorable and unfavorable attitudes expressed, the decision about whether or not to include a filter can have important substantive implications.

One set of relevant research has examined whether no-opinion filters enhance validity in the sense of detecting more true non-opinions. This work has focused on the fact that many respondents offer attitudes toward objects that are completely fictitious when asked unfiltered questions. Not surprisingly, including a no-opinion filter does dramatically reduce the number of respondents reporting such attitudes, though filters do not completely eliminate them. This suggests that filtering may enhance validity.

However, other evidence suggests that filtered questions are not superior to unfiltered questions in terms of data quality. First, studies examining test-retest response consistency indicate that reliability does not increase notably when filters are included. Second, associations between attitude reports and other variables do not seem to increase when filters are included. Third, it appears that respondents are especially attracted to no-opinion responses when selecting a substantive response would be especially cognitively burdensome. Therefore, it seems that no-opinion filters may not be desirable in attitude questions.

Response Sets and Response Styles

Response sets refer to the tendency for an individual to respond to questions in a particular fashion as a result of the structural features of the questions or the data-gathering situation, independent of the content of the questions. In contrast, response styles are response tendencies independent of content that are a function of dispositions of individual respondents, rather than a function of situational factors. A vast literature has investigated response styles and sets, most notably acquiescence and social desirability bias *(see* Paulhus, 1991). ·

Acquiescence refers to a tendency to agree with any item, irrespective of its content. This bias often occurs in items using "agree/ disagree" or "yes/no" formats. This tendency appears to be a result of both situational demands (e.g., the difference in social status between interviewers and respondents) and PERSONALITY characteristics (e.g., deferential personality). Some researchers believe that acquiescence can be eliminated by administering a large number of items, half of which express positive attitudes and half of which express negative attitudes. Combining across such a set of items may unconfound acquiescence with substantive responses in some cases. However, recent research suggests that this approach may often be unsuccessful, because different items can stimulate different levels of acquiescence, so counterbalancing these levels may be difficult. Fortunately, though, it appears that acquiescence bias can be 'eliminated by abandoning agree/ disagree or yes/ no questions and instead using forced choice formats that ask respondents to select one of two opposing substantive points of view.

Social desirability bias refers to a tendency to respond to questions in a way that is socially desirable. This bias can take two forms. In some cases, self-deception is involved, where people are inclined to perceive themselves in inaccurate and socially desirable ways. Alternatively, IMPRESSION MANAGEMENT motives can lead people to distort their presentations of themselves to others in socially desirable directions. Social desirability bias does indeed appear frequently in studies that have looked for it, and its magnitude appears to be a function of both situational and dispositional factors. The impact of social desirability bias may be reduced by conducting pretests to equate response alternatives in terms of their social desirability or to select items that are relatively unaffected by desirability. Additionally, scales measuring social desirability can be used to statistically remove the influence of this bias from other scales. In the case of impression management, assuring anonymity of responses can reduce social desirability biases.

Conclusion

Although elaborate procedures for attitude measurement were the norm in the early stages of empirical research in this area, practical considerations have led current researchers to adopt simpler, single-item approaches. However, this move ha$ increased the significance of the structural characteristics of the items used in terms of their potential impact on substantive findings. Fortunately, however, the accumulating body of literature we have reviewed provides numerous insights that aid researchers in designing their measures to tap attitudes as effectively as possible. With continued work in this area will come even clearer recommendations for optimal procedures.

8. Answer the following, each in not more than 250 words: **20x3=60**

(a) What is meant by effective communication? Evaluate various sources of effective communication.

TOPIC: *Language and Communication*
SUBTOPIC: *Effective Communication Training*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

Effective communication helps us better understand a person or situation, enables us to resolve differences, build trust and respect, and create environments where creative ideas, problem solving, affection, and caring can flourish. As simple as communication seems, much of what we try to communicate–and others try to communicate to us–gets misunderstood, which can cause conflict and frustration in personal and

professional relationships. By learning these effective communication skills, you can better connect with your spouse, kids, friends, and coworkers.

What is effective communication?

In the information age, we have to send, receive, and process huge numbers of messages every day. But effective communication is about more than just exchanging information. Effective communication requires you to also understand the emotion behind the information. It can improve relationships at home, work, and in social situations by deepening your connections to others and improving teamwork, decision-making, caring, and problem solving. It enables you to communicate even negative or difficult messages without creating conflict or destroying trust. Effective communication combines a set of skills including nonverbal communication, attentive listening, the ability to manage stress in the moment, and the capacity to recognize and understand your own emotions and those of the person you're communicating with.

While effective communication is a learned skill, it is more effective when it's spontaneous rather than formulaic. A speech that is read, for example, rarely has the same impact as a speech that's delivered (or appears to be delivered) spontaneously. Of course, it takes time and effort to develop these skills and become an effective communicator. The more effort and practice you put in, the more instinctive and spontaneous your communication skills will become.

Effective communication skills #1: Listening

Listening is one of the most important aspects of effective communication. Successful listening means not just understanding the words or the information being communicated, but also understanding how the speaker feels about what they're communicating.

Effective listening can:

- Make the speaker feel heard and understood, which can help build a stronger, deeper connection between you.
- Create an environment where everyone feels safe to express ideas, opinions, and feelings, or plan and problem solve in creative ways.
- Save time by helping clarify information, avoid conflicts and misunderstandings.
- Relieve negative emotions. When emotions are running high, if the speaker feels that he or she has been truly heard, it can help to calm them down, relieve negative feelings, and allow for real understanding or problem solving to begin.

Tips for effective listening

- If your goal is to fully understand and connect with the other person, listening effectively will often come naturally. If it doesn't, you can remember the following tips. The more you practice them, the more satisfying and rewarding your interactions with others will become.
- Focus fully on the speaker, his or her body language, and other nonverbal cues. If you're daydreaming, checking text messages, or doodling, you're almost certain to miss nonverbal cues in the conversation. If you find it hard to concentrate on some speakers, try repeating their words over in your head—it'll reinforce their message and help you stay focused.
- Avoid interrupting or trying to redirect the conversation to your concerns, by saying something like, "If you think that's bad, let me tell you what happened to me." Listening is not the same as waiting for your turn to talk. You can't concentrate on what someone's saying if you're forming what you're going to say next. Often, the speaker can read your facial expressions and know that your mind's elsewhere.
- Avoid seeming judgmental. In order to communicate effectively with someone, you don't have to like them or agree with their ideas, values, or opinions. However, you do need to set aside your judgment and withhold blame and criticism in order to fully understand a person. The most difficult communication, when successfully executed, can lead to the most unlikely and profound connection with someone.

- Show your interest in what's being said. Nod occasionally, smile at the person, and make sure your posture is open and inviting. Encourage the speaker to continue with small verbal comments like "yes" or "uh huh."

Effective communication skills #2: Nonverbal communication

When we communicate things that we care about, we do so mainly using nonverbal signals. Wordless communication, or body language, includes facial expressions, body movement and gestures, eye contact, posture, the tone of your voice, and even your muscle tension and breathing. The way you look, listen, move, and react to another person tells them more about how you're feeling than words alone ever can.

Developing the ability to understand and use nonverbal communication can help you connect with others, express what you really mean, navigate challenging situations, and build better relationships at home and work.

You can enhance effective communication by using open body language—arms uncrossed, standing with an open stance or sitting on the edge of your seat, and maintaining eye contact with the person you're talking to.

You can also use body language to emphasize or enhance your verbal message—patting a friend on the back while complimenting him on his success, for example, or pounding your fists to underline your message.

Tips for improving how you read nonverbal communication

- Practice observing people in public places, such as a shopping mall, bus, train, café, restaurant, or even on a television chat show with the sound muted. Observing how others use body language can teach you how to better receive and use nonverbal signals when conversing with others. Notice how people act and react to each other. Try to guess what their relationship is, what they're talking about, and how each feels about what is being said.
- Be aware of individual differences. People from different countries and cultures tend to use different nonverbal communication gestures, so it's important to take age, culture, religion, gender, and emotional state into account when reading body language signals. An American teen, a grieving widow, and an Asian businessman, for example, are likely to use nonverbal signals differently.
- Look at nonverbal communication signals as a group. Don't read too much into a single gesture or nonverbal cue. Consider all of the nonverbal signals you receive, from eye contact to tone of voice and body language. Anyone can slip up occasionally and let eye contact slip, for example, or briefly cross their arms without meaning to. Consider the signals as a whole to get a better "read" on a person.

Tips for improving how deliver nonverbal communication

- Use nonverbal signals that match up with your words. Nonverbal communication should reinforce what is being said, not contradict it. If you say one thing, but your body language says something else, your listener will likely feel you're being dishonest. For example, you can't say "yes" while shaking your head no.
- Adjust your nonverbal signals according to the context. The tone of your voice, for example, should be different when you're addressing a child than when you're addressing a group of adults. Similarly, take into account the emotional state and cultural background of the person you're interacting with.
- Use body language to convey positive feelings even when you're not actually experiencing them. If you're nervous about a situation—a job interview, important presentation, or first date, for example—you can use positive body language to signal confidence, even though you're not feeling it. Instead of tentatively entering a room with your head down, eyes averted, and sliding into a chair, try standing tall with your shoulders back, smiling and maintaining eye contact, and

delivering a firm handshake. It will make you feel more self-confident and help to put the other person at ease.

Effective communication skills #3: Managing stress

In small doses, stress can help you perform under pressure. However, when stress becomes constant and overwhelming, it can hamper effective communication by disrupting your capacity to think clearly and creatively, and act appropriately. When you're stressed, you're more likely to misread other people, send confusing or off-putting nonverbal signals, and lapse into unhealthy knee-jerk patterns of behavior.

How many times have you felt stressed during a disagreement with your spouse, kids, boss, friends, or coworkers and then said or done something you later regretted? If you can quickly relieve stress and return to a calm state, you'll not only avoid such regrets, but in many cases you'll also help to calm the other person as well. It's only when you're in a calm, relaxed state that you'll be able to know whether the situation requires a response, or whether the other person's signals indicate it would be better to remain silent.

Quick stress relief for effective communication

When stress strikes, you can't always temper it by taking time out to meditate or go for a run, especially if you're in the middle of a meeting with your boss or an argument with your spouse, for example. By learning to quickly reduce stress in the moment, though, you can safely face any strong emotions you're experiencing, regulate your feelings, and behave appropriately. When you know how to maintain a relaxed, energized state of awareness—even when something upsetting happens—you can remain emotionally available and engaged.

To deal with stress during communication:

- Recognize when you're becoming stressed. Your body will let you know if you're stressed as you communicate. Are your muscles or your stomach tight and/or sore? Are your hands clenched? Is your breath shallow? Are you "forgetting" to breathe?
- Take a moment to calm down before deciding to continue a conversation or postpone it.
- Bring your senses to the rescue and quickly manage stress by taking a few deep breaths, clenching and relaxing muscles, or recalling a soothing, sensory-rich image, for example. The best way to rapidly and reliably relieve stress is through the senses: sight, sound, touch, taste, and smell. But each person responds differently to sensory input, so you need to find things that are soothing to you.
- Look for humor in the situation. When used appropriately, humor is a great way to relieve stress when communicating. When you or those around you start taking things too seriously, find a way to lighten the mood by sharing a joke or amusing story.
- Be willing to compromise. Sometimes, if you can both bend a little, you'll be able to find a happy middle ground that reduces the stress levels for everyone concerned. If you realize that the other person cares much more about something than you do, compromise may be easier for you and a good investment in the future of the relationship.
- Agree to disagree, if necessary, and take time away so everyone can calm down. Take a quick break and move away from the situation. Take a stroll outside if possible, or spend a few minutes meditating. Physical movement or finding a quiet place to regain your balance can quickly reduce stress.

Effective communication skills #4: Emotional awareness

Emotions play an important role in the way we communicate at home and work. It's the way you feel, more than the way you think, that motivates you to communicate or to make decisions. The way you react to emotionally-driven, nonverbal cues affects both how you understand other people and how they understand you. If you are out of touch with your feelings, and don't understand how you feel or why you feel that way, you'll have a hard time communicating your feelings and needs to others. This can result in

96

frustration, misunderstandings, and conflict. When you don't address what's really bothering you, you often become embroiled in petty squabbles instead—arguing with your spouse about how the towels should be hung, for example, or with a coworker about whose turn it is to restock the copier.

Emotional awareness provides you the tools for understanding both yourself and other people, and the real messages they are communicating to you. Although knowing your own feelings may seem simple, many people ignore or try to sedate strong emotions like anger, sadness, and fear. But your ability to communicate depends on being connected to these feelings. If you're afraid of strong emotions or if you insist on communicating only on a rational level, it will impair your ability to fully understand others, creatively problem solve, resolve conflicts, or build an affectionate connection with someone.

How emotional awareness can improve effective communication

Emotional awareness—consciousness of your moment-to-moment emotional experience—and the ability to manage all of your feelings appropriately is the basis for effective communication.

Emotional awareness helps you:

- Understand and empathize with what is really troubling other people.
- Understand yourself, including what's really troubling you and what you really want.
- Stay motivated to understand and empathize with the person you're interacting with, even if you don't like them or their message.
- Communicate clearly and effectively, even when delivering negative messages.
- Build strong, trusting, and rewarding relationships, think creatively, solve problems, and resolve conflicts.
- Effective communication requires both thinking and feeling
- When emotional awareness is strongly developed, you'll know what you're feeling without having to think about it—and you'll be able to use these emotional cues to understand what someone is really communicating to you and act accordingly. The goal of effective communication is to find a healthy balance between your intellect and your emotions, between thinking and feeling.

Emotional awareness is a skill you can learn

Emotional awareness is a skill that with patience and practice can be learned at any time of life. You can develop emotional awareness by learning how to get in touch with difficult emotions and manage uncomfortable feelings, including anger, sadness, fear, disgust, surprise, and joy. When you know how to do this, you can remain in control of your emotions and behavior, even in very challenging situations, and communicate more clearly and effectively.

(b) Critically evaluate the relationship between intelligence and creativity. Cite experimental evidence in support of your answer.

TOPIC: *Thinking and Problem Solving*
SUBTOPIC: *Creative Thinking and Fostering Creativity*
LEVEL: *Medium*
NATURE: *Fundamental*

REFERENCE:

Creativity and intelligence are among the most important individual characteristics in a progressive society. The greatest advances and discoveries usually result from highly creative and intelligent people. Although the constructs of creativity and intelligence have been defined in many ways, creativity can be defined as "the ability to produce something that is both novel and appropriate" (Preckel, Holling, & Wiese, 2005, p. 160). Torrance (1969, as cited in Bracken, 1991, p. 174) suggested that, "Creative behavior occurs in the process of becoming sensitive to or aware of problems, deficiencies, gaps in knowledge, missing elements, disharmonies, and so on." Intelligence is commonly understood as the measurement I.Q.

The Relationship between Creativity and Intelligence

Researchers have long debated the nature of the relation between creativity and intelligence. Some have said that creativity and intelligence are two completely different and independent constructs, an assumption underlying earlier research in the field (e.g., Getzels & Jackson, 1962). Others have said that they are distinct but related constructs (Sternberg, 2000). Hayes (1989, cited in Sternberg, 2000) states that creativity and intelligence are not fundamentally related but that intelligence may be needed in order to display creativity. He noted, for example, that occupations with a type of creative freedom are more likely to require a higher level of education. Torrance (cited in Sternberg & O'Hara, 2000) concluded from his studies that intelligence and creativity are moderately related at best. Creativity has been said to be a subset of intelligence. Guilford (1950, 1967, 1970, 1975, cited in Sternberg, 2000), theorized that intellect is comprised of 120 different factors, with divergent production being the most relevant to creativity. Divergent production is the generalized gathering of knowledge and use of this knowledge to produce many different ideas in response to problems.

In contrast, Sternberg and Lubart's (1995) theory, as cited in Sternberg (2000), considers intelligence to be a subset of creativity. They posit that creativity is comprised of six different elements: intelligence, knowledge, thinking styles, personality, motivation, and the environment. Intelligence is comprised of three constructs: synthetic, analytical, and practical abilities. Synthetic ability is the ability to create different ideas that are appropriate to the task at hand. Analytical ability is the ability to give a judgment of the worth of one's ideas, which is measured using conventional intelligence testing. The last construct of intelligence is also relevant to creativity: practical ability is the ability to apply the intellectual skills gained to daily activities. All of these specific abilities are needed in order to exhibit creativity.

Other theorists have said that intelligence and creativity are related up to a certain IQ level and then unrelated among individuals with very high IQs (Preckel, Holling, & Wiese 2005). This hypothesis, known as the threshold hypothesis, states that creativity and intelligence are positively correlated up to a certain cut-off point (e.g., 120), beyond which there is no significant correlation.

Empirical results suggest there is at least some relatedness of creativity and IQ. For example, a meta-analysis of 100 studies investigating the threshold hypothesis found that there was a small, positive correlation (Kim, 2005). However, the relationship of creativity and intelligence may depend upon exactly what type of creativity is being measured and exactly what type of intelligence is being measured. Sligh, Conners, and Roskos-Ewoldsen (2005) used two different types of creativity measures as well as a composite IQ score to test the threshold hypothesis. The findings of the study showed that neither fluid intelligence (the ability to verbally or nonverbally reason abstractly) nor crystallized intelligence (knowledge gained from formal education and cultural experiences, verbal or nonverbal) showed support for the threshold hypothesis. Sligh et al. found that two specific aspects of creativity, generation and interpretation, were correlated differently with the composite IQ measure. The study suggests that looking into the different aspects of both intelligence and creativity could be helpful in understanding both constructs.

The empirical literature has extremely little to say about the relation between creativity and intelligence in the very young or the very old. Many studies involving older adults focus on either creativity or intelligence separately as they affect daily life, such as emotions and relationships. Ruth and Birren (1985) conducted a study that compared the creativity of three different age groups of adults from age 25 to 75. There was a decrease in creativity as well as in specific parts of intelligence in the sample as the ages of participants increased. Fuchs-Beauchamp, Karnes, and Johnson (1993) looked at the relationship between creativity and intelligence in children identified as eligible for admission to a program for gifted and talented students. Strength of this study includes the large sample of preschool students and the number of participants with an IQ over 120. The results showed moderate positive correlations between the different subscales of creativity and intelligence in children with an IQ less than 120. There was a decrease in correlation coefficients in the group of children with an IQ greater than 120. These results are consistent with the threshold hypothesis. However, due to the number of students with extremely high IQs the study cannot be generalized to the total preschool population.

Contemporary creativity research views intelligence and creativity as distinct traits that are only modestly related (for reviews see Batey & Furnham, 2006; Kauman, 2009; Kim, Cramond, & VanTassel-Baska, 2010; Runco, 2007). In his seminal work, Guilford (1957; Wilson, Guilford, & Christensen, 1953) took a more subtle view of the two abilities. He theorized that creativity is made up of a number of intellectual factors, with an emphasis on factors such as fluency, flexibility, and originality, that amount to a person's productive thinking ability. Guilford coined the now-famous distinction between convergent and divergent thinking—processes that lead to single, correct conclusions versus processes that generate multiple possibilities—but he viewed both as guided by similar traits and placed each within the umbrella of his Structure of Intellect model of cognitive abilities.

Later researchers, however, drew a sharper contrast between convergent and divergent thinking, equating the first with executive aspects of cognition and the second with associationistic aspects. Wallach and Kogan (1965), in their classic research, argued that divergent thinking reflects an associative process in which obvious, accessible ideas cue connected ideas, which in turn cue connected ideas, and so on. Creative ideas occur when the associative spread reaches distal concepts that are remotely related to the original concept. This explanation for divergent thinking is structural rather than executive: creative people are thought to have many loosely-related concepts, so their associative processes are more likely to generate remote, distal concepts. Wallach and Kogan's model was greatly influenced by Mednick's (1962) model of creative thought—a similar structural model—that explained individual differences in creativity with differences in whether "associative hierarchies" were steep or flat. People with flat hierarchies—i.e., with numerous and loose conceptual connections—were potentially more creative.

Since Wallach and Kogan's work, creativity researchers have continued to explain individual differences in creativity in terms of differences in the structure of knowledge, particularly the nature and extent of associations between concepts (see reviews by Runco, 2007 and Weisberg, 2006). For this reason, it is not surprising that modern creativity research expects weak effects of fluid and executive processes on creative thought, given that creative ideation is viewed as structural and associationistic. Moreover, a large body of work supports Wallach and Kogan's view that intelligence and creativity are modestly related. In a recent meta-analysis of 447 effect sizes, Kim (2005) found an average weighted effect size of $r=.174$.

An executive interpretation of divergent thinking tasks Despite the historical emphasis on knowledge structure and associative processes, recent work has reinterpreted divergent thinking tasks as essentially executive tasks that are founded on managing interference. Gilhooly et al. (2007) recently proposed that generating unusual uses for a common object entails managing interference from several sources: the object's obvious, highly accessible uses (e.g., using a brick to build a wall or make a path); concrete features of the object, which interfere with abstracting general features that can transfer across domains (e.g., getting stuck on a brick's shape and density); and interference from one's own previous responses, which often intrude into the list of responses. To generate creative responses, people have to identify a useful strategy and use it in the face of this interference, so executive cognition is central to successful performance on a divergent thinking task.

Two studies provided support for an executive interpretation of creative ideation. In their first study, Gilhooly et al. (2007) used verbal protocols—also known as think-alouds—to identify the strategies people spontaneously used for an unusual uses task, and they rated each task response for novelty. They found several common strategies that didn't foster novel responses, such as memory retrieval (e.g., scanning long-term memory for examples of creative uses) and self-cueing (e.g., repeating the name and features of the object). But several infrequent and abstract strategies were associated with responses of a more creative nature, such as disassembly (e.g., breaking the object into parts). In a second study, a measure of executive ability (a letter fluency task) significantly predicted coming up with more "new" responses (i.e., generated during the task itself) but not "old" responses (i.e., retrieved from long-term memory).

The creativity-and-intelligence controversy

In the psychology of creativity, most reviews of the creativity-and-intelligence controversy have concluded that creativity and intelligence are distinct abilities with minor overlap (e.g., Batey & Furnham, 2006;

Kaufman & Plucker, 2011; Kim, Cramond, & VanTassel-Baska, 2010; Runco, 2007). Since Wallach and Kogan's (1965) landmark work on this topic, research has typically found that creative cognition—usually measured with divergent thinking tasks—covaries modestly with intelligence. A recent meta-analysis of the relationship between intelligence and divergent thinking found an overall effect of r=.17 (Kim, 2005).

At the same time, many contemporary researchers have found that there are good reasons to expect stronger relationships between intelligence and creative cognition. Generating creative ideas—ideas that are both novel and appropriate to the purpose at hand—requires identifying and implementing strategies for idea generation (Gilhooly, Fioratou, Anthony, & Wynn, 2007; Nusbaum & Silvia, 2011a), exerting control over attention and thought (Vartanian, 2009; Zabelina & Robinson, 2010; Zabelina, Robinson, Council, & Bresin, in press), making decisions and refining initial ideas (Finke, Ward, & Smith, 1992; Gabora, 2005; Vartanian, 2011), and inhibiting obvious and inapt ideas (Nusbaum & Silvia, 2011a).

If this view of creative cognition is right, then fluid and executive abilities should be central to the creative process. But past reviews and Kim's (2005) meta-analysis conclude otherwise, so an executive interpretation of creative thought is understandably controversial.

(c) How would you measure emotion of an individual? Discuss the effect of emotion on behavior.

TOPIC: *Motivation and Emotion*
SUBTOPIC: *Measurement of Motivation and Emotion*
LEVEL: *Easy*
NATURE: *Fundamental*

REFERENCE:

METHODS FOR MEASURING EMOTION

As mentioned earlier, the components of emotion can be conveniently categorized into three broad, only somewhat overlapping, output systems (after Lang, 1979): (a) language, which includes evaluative self-reports of experience, modifications of speech patterns by emotions, and expressive communications; (b) behavior, which includes overt acts and functional behavior sequences, such as flight or fight, approach, and threat behavior, as well as the modulation of other behaviors by emotion; and (c) physiology, which includes central and peripheral nervous system changes that support emotional behavior or prepare the person for responding, as well as somatic expressive changes and hormonal changes that coordinate the discharge and replacement of energy stores. As presented by Bradley and Lang (2000), these three output systems provide a three-dimensional framework for considering measures of emotion. We present this framework in the form of a data box (Cattell, 1988; Larsen, 1989).

Emotions can be thought of as events that produce variability along these three main dimensions. Each dimension of this framework refers to a category of variables that have been theoretically and empirically related to emotion. A theoretical challenge in the emotion area is to sort out this response diversity by explaining when response patterns across the three dimensions will be integrated and when they will be discordant. Until then, researchers would be well advised to include, at the least, measures from each of the three dimensions.

We turn now to a brief review of various measures that make up each of the three dimensions. Our intent is to provide a sampling of the major methods that have been developed in the three broad areas of emotion output. We will not provide an exhaustive review of what is known about each measure, nor is our review intended to be exhaustive within each type of measurement method. Rather, our intent is to provide examples of specific methods, describe a few strengths and weakness of each, and provide references for interested readers to pursue for further details.

Emotion Assessed Through Language

Self-report evaluations of emotion. Self-report measures, where participants provide an evaluation of their emotional experience, form the most diverse yet most widely used set of assessment tools for measuring emotion (Larsen & Fredrickson, 1999). Measures range from rating scales and adjective checklists, to analog scales and real-time rating dials. Proponents of self-report measures (e.g., Baldwin, 2000) assume that participants are in a privileged position to monitor, assess, and integrate information about their own emotions, and therefore self-report measures should not be thought of as second-rate proxies for better measures. Critics of self-report measures (e.g., Schaeffer, 2000), on the other hand, argue that there are so many biases, distortions, and methodological limitations that self-reports of anything, even one's home address, are fraught with error and misinformation. Self-report measures are nevertheless the most widely used assessment tools in emotion research.

Although there are a great many self-report instruments, considerable similarities can be found among them. Here we present a few exemplars and highlight themes and issues common to self-report measures. Reviews of specific self-report instruments can be found in Larsen and Fredrickson (1999), MacKay (1980), and Stone (1995; Stone, Turkkan, Bachrach, Jobe, Kurtzman, & Cain, 2000).

An assessment strategy with a good deal of face validity is simply to ask participants to rate how they are or were feeling on a single emotion dimension. That dimension might be a global affective evaluation (e.g., How unpleasant are you feeling?) or a specific emotion (How angry do you feel?). And the response scale might be unipolar (not at all angry to extremely angry) or bipolar (unpleasant to pleasant), with response options that are Likert type scales (e.g., 5-, 7-, or 9-point formats). Or the response might be in a checklist format, where the respondent indicates whether or not a specific emotion was experienced. Such measures are simple to construct, easily understood by participants, and brief to administer. Virtually any emotion term can anchor a scale or be put onto a checklist, making self-report indispensable for researchers targeting specific, discrete emotions, as well as those researchers using multiple items to reflect global dimensions of emotion.

A variation on self-report is the experience sampling method, where participants make frequent reports over an extended time period. Although this method allows researchers to ask unique theoretical questions about emotion (e.g., Larsen, 1987), the measurement concerns remain mainly those associated with simple self-report. See Bolger, Davis, and Rafaeli (2003) for a review of this method.

A variation on rating scales makes the response a visual analog that presents the participant with a horizontal line separating two opposing adjectives, which lessens stereotyped responding. A related technique is to make the question itself an analog of the emotion being assessed. For example, the participant might be presented with a series of five cartoon faces, going from a neutral expression on one face to an extreme frown on another. This has the advantages of being useful with participants for whom adjectives might not be meaningful, such as very young children or participants from different language groups.

Another useful strategy in self-report is to have the participants indicate, in real time, how they are feeling by turning a dial, moving a mouse, adjusting a computer display, or in some way modifying an analog display of the emotion on which they are reporting. The general strategy across these techniques is to collect self-reports of subjective experience on a moment-by-moment basis, either online as the emotion is experienced, or retrospectively as the original episode is "replayed."

Conceptually, the most basic real-time self-report measure can be viewed as a single-item measure with a temporal dimension added. Using some mechanical input device (e.g., a mouse or joystick), respondents adjust a computer display as often as necessary so that it always reflects how they are feeling each moment throughout an extended episode (e.g., Schuldberg & Gottlieb, 2002, used such a device to obtain 1,400 affect readings over 2.5 minutes for each subject). Several researchers have described continuous "rating dials" of this sort (Bradley & Lang, 2000; Bunce, Larsen, & Cruz, 1993; Fredrickson & Kahneman, 1993; Gottman & Levenson, 1985). Like rating scales more generally, rating dials may use either bipolar (very negative to very positive) or unipolar verbal anchors (no sadness at all to extreme sadness) and either Likert type or visual analog scales.

Advantages to these procedures include automating self-report data collection, the ability to calibrate self-reports with other emotion measures (e.g., physiology, facial expressions) in the temporal stream, and the ability to use the technique "off line" to have participants continuously, though retrospectively, report on the emotions they were experiencing (e.g., Gottman & Levenson, 1985; Levenson & Gottman, 1983). The major disadvantage is the need for specialized equipment and the fact that the participant's attention is partially focused on the rating device. Moreover, it seems likely that continuously monitoring one's emotions may lead to a form of fatigue or may be so intrusive that it actually alters the respondent's emotions. Another drawback of this assessment strategy is that the techniques are limited to the self-report of just one or two dimensions. Although it is technically feasible, for example, to create a whole bank of rating dials (e.g., anger, fear, sadness, disgust, attraction, enjoyment, contentment), a limiting factor would be the respondent's ability to track the ebb and flow of multiple discrete emotions simultaneously.

Another category of self-report measures consists of the many standardized multi-item emotion inventories. Some of these inventories are checklists, whereas others are rating scales. These instruments are essentially variations on the self-report themes mentioned earlier, with differences having to do primarily with response scales, the number and nature of the emotion adjectives, the scoring and scale names, and the instructions that accompany the self-report tasks. The advantages of these inventories include their theory- or statistically guided development, empirical refinement and standardization, the development of norms (which allow cross-study comparisons and even meta-analysis; e.g., Larsen & Sinnett, 1991), and the accrual of research findings on specific measures and specific constructs-measure units.

One of the first self-report emotion inventories formally constructed was the 130-item Mood Adjective Checklist (MACL; Nowlis & Green, 1957). Not literally a checklist, the instructions ask the participant to rate how they feel on a Likert scale. Scoring results in 12 factor scores: aggression, anxiety, urgency, elation, concentration, fatigue, social affection, sadness, skepticism, egotism, vigor, and nonchalance. Other researchers have proposed a simpler positive-negative valence scoring scheme (Stone, 1981). The MACL has not become widely used, most likely because it was never formally published (the original version appeared in an unpublished Naval Technical Report, Nowlis & Green, 1957).

A self-report emotion measure that eclipsed the MACL is Zuckerman and Lubin's (1965) Multiple Affect Adjective Checklist (MAACL). It is very similar to the MACL in length, with the MAACL having 132 items. The majority of the items overlap between the two inventories. The MAACL has become the most widely used self-report emotion assessment instrument in the psychological literature (Larsen & Sinnett, 1991). The MAACLs success is likely due to the fact that it is distributed by a professional test publisher and comes with a user manual, annotated references, developmental history, and psychometric properties, along with scoring keys and answer sheets. Other reasons for its popularity might be the checklist format, which makes administering the MAACL much faster than the MACL. And finally, the MAACL has only 3 subscales (depression, anxiety, and hostility), compared to 12 on the MACL. In 1985 Zuckerman and Lubin published a revised version of the Multiple Affect Adjective Checklist (MAACL-R). The revision mainly concerns the scoring format, which now allows for several pleasant emotion scores as well as global positive and negative affect and sensation seeking.

This is a good point to mention the issue of response formats. The MAACL and its revision are in the form of checklists, in which the subject merely indicates the presence or absence of a particular emotion by checking a box. Some researchers have argued that checklists are particularly susceptible to response styles and other forms of non-random error. Bentler (1969) argued against using checklists in psychometric assessment. Green, Goldman, and Salovey (1993) demonstrated that checklist emotion assessments contain significant non-random error, and they advised caution when analyzing or interpreting checklist data. However, more recently, Schimmack, Bockenholt, and Reisenzein (2002) demonstrated that checklist and Likert scale affect self-reports yield very similar covariance structures. The question of the impact of response format on affect ratings remains open.

Although several rating scales are available (see Stone, 1995), one of the more recent introductions is the Positive Affect Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). The PANAS is based on a dimensional model of emotion, in particular the circumplex model (Russell, 1980; Watson & Tellegen,

102

©VISION IAS
www.visionias.wordpress.com
Contact: +91-9650617807, +91-9968029039

©Numerons
www.numerons.com
Contact: +91-9560344245

1985). Of the eight potential scores derivable from the circumplex model (Larsen & Diener, 1992), the PANAS focuses on two of these: Positive Affect (PA; high arousal pleasant), and Negative Affect (NA; high arousal unpleasant). The PANAS contains 10 items on each of the two scales. The items are mood adjectives and are rated on a 5-point scale, labeled as "not at all or slight," "a little," "moderately," "quite a bit," and "very much." The PA and NA scales were constructed to be uncorrelated, and they generally are (though see Zautra, Berkhof, & Nicolson, 2002, for exceptions).

Like most self-report measures, research on the validity of the PANAS has been primarily correlational. For example, extraversion correlates with frequent reports of PA, and neuroticism correlates with frequent reports of NA. In one of the first experimental studies of the PANAS, Larsen and Ketelaar (1991) induced emotions in the laboratory using guided imagery. They found that the positive induction increased PA but did not lower NA, and the negative induction increased NA but did not lower PA. Similar experimental findings on the independence of PA and NA under different inductions, using naturalistic success and failure feedback on exam performance in college students, were found by Goldstein and Strube (1994). This differential sensitivity to positive and negative emotion inductions supports the construct validity of the PANAS. Nevertheless, researchers should be very clear that the PANAS does not measure discrete emotions, which other scales do. The PANAS has its greatest utility in the assessment of the broad emotion dimensions of PA and NA.

Evaluation of self-report methods. Self-report methods are perhaps the most efficient techniques for measuring emotions. Nevertheless, they rely on assumptions that research participants are both able and willing to observe and report on their own emotions. Some issues concern a person's ability to self-report their emotions. Self-report requires memory, either working memory or longer-term memory, and so a variety of memory distortions may compromise a report (Feldman Barrett, 1997). Self-report also requires the perception of something on which to report. It is possible that a person may "have" an emotion in a nonverbal channel (e.g., autonomic activation or action tendency) yet never label that experience and hence not perceive it as an emotion at all (Tranel & Damasio, 1985). Moreover, some persons may repress emotional experiences, particularly negative or inappropriate emotional experiences, resulting in biased or incomplete report of emotions (Cutler, Bunce, & Larsen, 1996). Certain populations, for various reasons, may have meager or inaccurate comprehension of verbal information. For example, cultural psychologists have argued that some cultures have emotions, or emotion terms, that are not identifiable in other cultures (e.g., Mesquita & Frijda, 1992).

Regarding the second assumption—that participants must be willing to report on their emotions—the issue here is mainly one of response sets, where responses to items might be based, not on the emotion content of the items, but on some other factor, such as their social desirability. Here the participant is responding to the items in a manner that creates a positive impression. A different response set is extreme responding, where a participant may be motivated to use scale endpoints or large numbers in describing their emotions, a response set that can greatly distort the covariance structure of a set of ratings (Bentler, 1969).

Another potential problem with self-report is measurement reactivity, where the actual process of measurement alters the psychological construct being measured. Administering an emotion self-report may, in fact, influence the emotional state of interest. Another issue arises when researchers want to assess emotion two or more times, as in within subject experimental designs or in experience sampling studies of emotion. One potential effect of repeated emotion measurement is stereotypic responding (Stone, 1995), where participants settle into a response profile that does not change much across the assessment occasions.

Self-report emotion measures require that subjects engage in a number of psychological processes to arrive at a rating. Understanding these processes has both theoretical as well as measurement implications. For instance, providing a global self-report implicates memory processes, as respondents recall the targeted episode, as well as aggregation processes, as respondents in some manner integrate their multiple and often varied momentary experiences into an overall rating. Both of these mental processes may obscure or misrepresent dynamic changes in emotion as experienced over time. For instance, Kahneman and his colleagues have documented that people's global reports of pain episodes draw highly from the momentary

affect experienced at the most intense point during the episode, as well as the final moments of the episode, with the duration of the emotional experience largely neglected in the global self-report (Fredrickson & Kahneman, 1993; Kahneman, 1999; Kahneman, Fredrickson, Schrieber, & Redelmeier, 1993; see also Thomas & Diener, 1990 for related issues).

Other language parameters related to emotion. Another language-related channel with potential as a measure of emotion is the voice. Vocalization may be sensitive to emotion-related changes in the body (e.g., muscle tension, respiration rate, and blood pressure). Vocal analysis for emotion has traditionally followed one of two possible strategies. The simplest strategy is to have humans listen to audio-taped speech and evaluate the speaker's affective state. A more technologically advanced strategy is to have audiotapes digitized and analyzed by computer.

The ability of untrained listeners to correctly recognize or infer speakers' emotional states has been evaluated in several studies (e.g., Scherer, 1986; Scherer, Banse, Wallbott, & Goldbeck, 1991; van Bezooijen, 1984). In these studies actors are used to produce sentences in a way that imparts a specific emotional tone (e.g., anger, fear, disgust, joy, sadness). The speech samples are then stripped of vocal content and are then played for naive listeners who judge which emotion they perceive in the vocalization. Correct selection rates across these studies average around 55%, a rate four to five times what would be expected by chance (Pittam & Scherer, 1993). Some emotions are more easily recognized by naive raters than others: Sadness and anger are best recognized, whereas disgust, contempt, and joy are least recognized in vocalization samples (Pittam & Scherer, 1993; van Bozooijen, Otto, & Heenan, 1983).

Studies also suggest that arousal level may be better transmitted by vocal cues than is specific hedonic content (i.e., Apple & Hecht, 1982; van Bozooijen et al., 1983). Reviews of recent research suggest that although perceivers are more accurate in judging nonspecific arousal from vocal parameters, they are nevertheless well above chance in judging pleasantness and specific emotions from speech samples that have had the verbal content removed (Bachorowski, 1999). A particularly impressive set of results is reported by Scherer, Banse, and Wallbott (2001). These researchers used professional German actors to produce vocal samples spoken in fear, anger, sadness, joy, and neutral vocal tones. The actual verbal content was then stripped away, leaving only vocalization. The samples were then taken to nine different countries in North America, Asia, and Europe, where participants from different language groups listened to the vocalizations and rated the likely emotions. Overall accuracy averaged 66%, a figure well above chance.

Researchers studying digital voice analysis are still searching for the parameters that best reflect emotion. Parameters typically assessed are (a) fundamental frequency, perceived as overall voice pitch; (b) small perturbations in the fundamental frequency; (c) intensity, indexed in decibels; and (d) speech rate or tempo (Scherer, 1986). Whereas acoustical analysis of speech most accurately reflects the nonspecific arousal of the speaker (Bachorowski & Owren, 1995), it falls far short of identifying specific emotions. For example, positive and negative emotional states are often not reliably distinguished with acoustical parameter (Scherer, 1986; see also Pittam & Scherer, 1993). Because untrained listeners can distinguish specific emotions from voice samples, there must be some acoustical cues for affect. However, at this time, researchers are still searching for those cues. See Russell, Bachorowski, and Fernandez-Dols (2003) for a recent review of vocal measures of emotion.

Emotion Assessed Through Behavior

Behaviors that are linked to emotions range from the very simple, such as defensive reflex actions, to the complex, such as sequences of action tendencies. Emotions likely evolved to produce adaptive actions, such as to approach desired objects or to withdraw from dangerous objects, as well as to support more flexible action tendencies associated with survival. Researchers may take advantage of these behavioral outputs to estimate emotions.

Behavior action tendencies. One behavioral manifestation of emotion concerns the action tendencies that become more or less likely during emotion. Tasks that inquire about various actions or intentions may be linked to emotion states. One task is to ask participants how much they would like to engage in various

behaviors, such as talk with a good friend, engage in some exercise, or have a pleasant meal. Teasdale and colleagues (Teasdale, Taylor, & Fogarty, 1980) reported that this task is sensitive to depressed mood or sadness, which has the action tendency of social withdrawal. When sad, people often lose interest in activities that formerly gave them pleasure. Sadness is also thought to be associated with depressed psychomotor function. Writing speed, for example, is negatively correlated with sadness and depression (see Velten, 1968, who used this task as one criterion measure in the validation study of the mood induction that bears his name). Other psychomotor tasks that have been used in emotion research include letter cancellation and smooth pursuit motor tracking tasks. Performance speed is most sensitive to sadness or depressed emotional states. Pleasant emotions, however, do not appear to increase psychomotor speed.

A variety of other behavioral tasks have been shown to be sensitive to affective states (Mayer, 1986; Mayer & Bremer, 1985; Mayer, Mamberg, & Volanth, 1988). One category of emotion-sensitive tasks consists of judgment tasks. One assessment strategy is to have participants make probability estimates of the likelihood of various good and bad events. For example, participants may be asked the probability of being killed in an airplane crash, dying in a car accident, or contracting cancer in their lifetime. It has been shown that persons in generalized unpleasant emotional states overestimate the probability of such bad events (Johnson & Tversky, 1983). The converse—increased probability estimates of good events while in positive emotional states—also appears true (Zelenski & Larsen, 2002). Results also apply to specific emotions; for example, fearful people make pessimistic judgments of future events (Lerner & Keltner, 2000). General appraisals of events also show emotion-specific patterns (Herrald & Tomaka, 2002; Siemer, 2001).

Another emotion-sensitive behavioral task is to ask participants to generate associations to positive, neutral, and negative stimuli. For example, have participants write down as many words as come to mind in 60 seconds when they hear each of the following stimulus words: happy, disappointed, generous, destroy, peace, or pain. Mayer and Bremer (1985) showed that performance on this task correlated with naturally occurring mood. Seidlitz and Diener (1993) used a variation wherein participants recalled as many happy experiences from their own life as they could in a given time period. Participants higher on trait-positive affect recalled more pleasant experiences, in the same time period, than participants lower on trait happiness. Teasdale and colleagues (Teasdale & Fogarty, 1979; Teasdale & Russell, 1983) have also demonstrated that emotion inductions influence recall of pleasant and unpleasant events in predictable (i.e., hedonically consistent) ways.

Another behavioral strategy for assessing emotion involves various information processing parameters. Reaction times in lexical decision tasks, for example, have been shown to be sensitive to affective states (Chains & Krane, 1988). For example, the participant's task might be to judge whether a string of letters represents a word or a non-word. On each trial the letters represent either: a non-word, an emotion word (e.g., anger), or a neutral word (e.g., house). Participants in positive affective states are quicker and sometimes more accurate at judging positive words compared to participants in neutral states, and vice versa for unpleasant moods (Niedenthal & Setterlund, 1994).

A related assessment task is to present participants with incomplete word stems and ask them to add letters to complete the word. Word stems are selected so that they can be completed as an emotion term or as a neutral term. For example, ANG_ could be completed as ANGER or as ANGLE or ANGEL or as ANGLO; JO_ could be completed as JOY or as JOB (e.g., Rusting & Larsen, 1998). A related technique is the use of homophones (words that sound alike but have different meanings). With this technique, the subject hears the word (die or dye, for example) and is asked to write that word. Participants in an unpleasant mood are more likely to write or complete the word stems in a manner congruent with their mood (Halberstadt, Niedenthal, & Kushner, 1995).

Behaviors that are enhanced or disrupted by emotion. So far we have discussed how certain behaviors directly follow from emotional states, and how specific emotion-related tasks may be influenced by the emotional state of the participant. However, other categories of behaviors, such as the defensive reflex or perception or attentional control, may be enhanced or disrupted by emotion, and thus might be used as an indicator or measure of emotion (Compton, 2000). One such emotion-sensitive task relies on a very simple behavior—the startle reflex. The startle reflex involves a rapid shutting of the eyes (blink), pulling the chin

down, and a rapid inhalation. The startle reflex is easy to elicit through the application of a sudden and loud acoustic stimulus. Startle potentiation refers to an increase in the startle response (measured as a faster or stronger eye blink) when the person is startled while they are in an unpleasant emotional state (Vrana, Spence, & Lang, 1988). The researcher most responsible for developing this technique in humans is Peter Lang (e.g., Lang, Bradley, & Cuthbert, 1990). Lang and colleagues, as well as others (Skolnick & Davidson, 2002), have demonstrated startle potentiation for unpleasant emotions compared to neutral states. The converse—slower and weaker startle when in positive emotional states—is rarely found.

A final behavioral paradigm with potential for measurement concerns the effects emotion has on cognitive parameters such as attention (Buodo, Sarlo, & Palomba, 2002). One effect is the automatic vigilance effect (Cothran, Zelenski, Prizmic, & Larsen, 2003; Pratto & John, 1991), which refers to the "grabbing" of attention by aversive or threatening information. The so-called emotional Stroop paradigm is one example of the automatic vigilance effect, where, in naming the colors of various words, people are generally slower to name the color if the word is threat-related. Presumably, threatening stimuli are processed more carefully, especially if one is already in an aversive emotional state, resulting in slowing on the primary, non-emotional task (color naming). Another behavioral paradigm, where cognitive parameters are influenced by both the emotional state of the participant, as well as the emotional content of the stimuli, concerns emotional priming (Wentura & Rothermund, 2003), as well as other irrelevant feature tasks, such as the Affective Simon task (De Houwer, 2003), where the participant in supposed to ignore the emotional content of a stimulus while responding to some other relevant feature.

Emotion Assessed Through Physiology

Emotion output that can be assessed with physiological methods can be divided into two categories: the somatic changes and changes reflecting autonomic or central nervous system activity. The somatic changes most useful to emotion researchers concern muscle movements associated with emotional expression, particularly those somatic changes on the face.

Measures of somatic change. One useful measurement strategy is to have an observer rate how much emotion a target participant appears to be feeling, based on expressive cues. The observers might be "experts" on the target person's emotional experiences (e.g., a spouse or a therapist). One limitation is that observer reports are based on social attributions of a target's emotional state, and such attributions will be limited by the information available, biased by a target's impression management strategies, or even influenced by the raters' own level of emotion being rated (Marcus & Miller, 1999). As such, observer ratings of emotion are probably best used in combination with other measures. One way to limit attributions is to use trained observers. A standardized training system for observers is the Specific Affect Coding System (SPAFF; Coan & Gottman, in press; Gottman & Krokoff, 1989; Gottman & Levenson, 1992; for a brief review, see Gottman, 1993). This system separates expressed emotion into specific categories of positive and negative categories (e.g., interest, affection, humor, validation, excitement/joy, anger, belligerence, domineering, contempt, disgust, tension, sadness, whining, and defensiveness). SPAFF training involves recognizing and attending to important facial, gestural, and vocal markers of emotion. Significant benefits of observer ratings are that they can be unobtrusive, can be used in naturalistic settings, are inexpensive and fast, and can provide emotion measures from a few visible cues.

Some somatic coding systems are based on specific observable changes in facial muscles. One such system for coding emotion in the face is the Facial Action Coding System (FACS; Ekman & Friesen, 1975, 1978). The FACS consists of 46 anatomically based "action units" (or AUs), which refer to a specific observable change in the face. For example, AU 1 raises the inner brows, AU 9 wrinkles the nose, and AU 12 raises the outer lip corners. The system requires extensive training and certification for reliable use (cf. Ekman & Friesen, 1975, 1978). A drawback of FACS is the extensive amount of time needed to code expressions. FACS scoring requires about 1 hour of coding for each minute of videotape (depending on the density of facial action). Researchers are developing computer vision to undertake the tiresome task of facial action coding. One of the more advanced systems is that being developed at Carnegie Mellon University under the guidance of Jeffry Cohn (see Cohn, Zlochower, Lien, & Kanade, 1999), which is able to accurately code approximately half of the FACS action units in real time. Alternatively, Ekman and others have developed more global

coding systems, which are based on fewer AUs, for coding facial action (e.g., EMFACS by Ekman & Friesen, see Fridland, Ekman, & Oster, 1986; MAX by Izard, 1979).

Somatic facial assessments may also be obtained using physiological measures of muscle contractions. The neural activation of the muscles produces action potentials that can be directly measured on the surface of the skin using electromyography (EMG) using two electrodes placed over the muscle of interest. The amount of electrical activity detected is directly proportional to the magnitude of contraction. Detailed descriptions of facial electromyographic technique may be found in Cacioppo, Petty, Losch, and Kim (1986). EMG is able to assess muscular contractions that are too small to produce visible changes (i.e., not FACS codable; Cacioppo et al, 1986). Such sensitivity has a disadvantage, however, in that electrical signals from sites other than the muscle of interest may also be detected during EMG assessments. Researchers interested in measuring emotions with facial EMG should have training in electrophysiological technique or collaborate with someone with such expertise.

Physiological measures of nervous system activity associated with emotion. Emotions are closely tied to tendencies to act in specific ways, and changes in the nervous system occur primarily to support these actions (Frijda, 1986; Lazarus, 1991). In terms of the autonomic nervous system (ANS), a few researchers hold the view that distinct emotions are associated with distinct ANS activity (e.g., Levenson, Ekman, & Friesen, 1990). Empirical support for specific autonomic patterns being associated with specific emotions has been obtained in several studies. However, the cumulative data on specific emotional "signatures" are mixed and therefore remain inconclusive (for reviews, see Cacioppo & Gardner, 1999; Cacioppo, Klein, Berntson, & Hatfield, 1993; Levenson, 1992; Zajonc & Mclntosh, 1992).

Diverse autonomic measures have been used to assess emotion, some more fruitfully than others. We will mention here only a couple of the more promising measures and advise the interested reader to consult Cacioppo, Tassinary, and Berntson (2000); Stern, Davis, and Ray (1992); or Hugdahl (1996) for more details. Electrodermal activity, especially skin conductance, is a widely accepted and reliable measure used in emotion research. Another category of measures is based on respiratory activity. Perhaps the largest category of measures is those based on cardiovascular activity. This last set includes measures such as heart rate, diastolic and systolic blood pressure, cardiac output, stroke volume, and total peripheral resistance. Readers interested in cardiac measures should consult Sherwood (1993) and Sherwood et al. (1990) for details on impedance cardiography. Other researchers assess the link between respiratory and cardiovascular activity (e.g., respiratory sinus arrhythmia or heart rate variability), which appears related to emotion state (Grossman, van Beek, & Wientjes, 1990; Porges, 1995). It should be noted that professional polygraphers typically employ a multi-method approach, using measures of skin conductance, respiration, and heart rate to infer the emotion of guilt.

Researchers have recently begun to refine central nervous system measures of emotion. Scalp-recorded brain electrical activity, or electroencephalogram (EEG), has been used successfully to distinguish pleasant and unpleasant emotion states (e.g., Schmidt & Trainor, 2001), as well as individual differences in affective style (for a review, see Davidson, 1993). Other more localized imaging measures of emotion-related changes in the brain are on the horizon as well, including functional MRI (for an overview, see Berthoz, Blair, Le Clec'h, & Martinet, 2002). The versatility of functional imaging methods for studying mechanisms of emotion is significant, given its superior spatial resolution (Mayberg & McGinnis, 2000). The temporal resolution is not as good as EEG measures, however.

Many practical issues emerge when contemplating the use of physiological measures. First, these measures are typically invasive. Some measures (e.g., being inserted into a large MRI magnet) might elicit emotions (e.g., panic) themselves. Less-invasive measures are pulse rate and skin conductance. Impedance cardiography uses metal bands that cir- cle a participant's neck and chest in several locations. Attaching these requires participants to partially disrobe. Blood pressure assessment typically uses pressurized cuffs that, when inflated, draw attention and sometimes even cause pain. Physiological measures also usually restrict participants' mobility because of wires that connect them to amplifiers and recording devices. Bodily movement can also create artifacts in measurement. In general, the use of physiological assessments requires special efforts on the part of both the researcher and the participants, but may potentially pay off with a unique methodological perspective or window on the emotional state under investigation.

EFFECTS OF EMOTIONS ON BEHAVIOR

Emotion may be an internal, purely subjective experience, but it also has an observable behavioral component. The links between inner feelings and outward expressions are numerous: We smile when we're happy, cry when we're sad, blush when we're embarrassed, stand tall when we feel proud, drag our feet when we're down, press our lips in anger, bow our heads in shame, and wrinkle our faces in disgust. These behavioral expressions of emotion serve two functions. First, they provide us with a means of *nonverbal communication.* People often use words to tell others how they're feeling. But by smiling, frowning, turning bright red in the face, shrugging the shoulders, or winking an eye, we also communicate our feelings nonverbally—which encourages others to approach us, or stay away. Thus, Alan Fridlund (1994) has argued that our expressive behaviors serve more as *signals* to other people than as *symptoms* of how we feel, making the display of emotion an inherently social experience. The second effect of behavioral expression is to provide us with *sensory feedback.* In 1872, Darwin theorized that the expressions that we make clarify and intensify emotional experiences by providing us with bodily feedback about how we feel. In short, the expressive component of emotion has two audiences: other people and ourselves.

NONVERBAL COMMUNICATION

Knowing how another person is feeling can be tricky because people sometimes try to hide their true emotions. Think about it. Have you ever had to suppress your rage at someone, mask your disappointment after failure, feign surprise, or pretend to like something just to be polite? Sometimes we come right out and tell people how we feel. But often we actively try to conceal our true feelings. In instances like these, observers tune in to a silent language—the language of nonverbal behavior.

Facial Expression

What kinds of nonverbal cues do people use to judge how someone is feeling? In *The Expression of the Emotions in Man and Animals,* Charles Darwin (1872) argued that the *face* communicates emotion in ways that are innate and are understood by people all over the world. Contemporary research provides strong support for this proposition. In a groundbreaking study, Ekman and Friesen (1974) showed thirty photographs like those on page 399 to participants from New York to New Guinea—including Argentina, Borneo, Brazil, and Japan—and asked them to guess the emotion being portrayed in each photo. The results of this study, and of many others like it, indicate that people can reliably identify six emotions: joy, fear, anger, sadness, surprise, and disgust. In the experiment described at the start of this chapter, people from ten different countries exhibited high levels of agreement in their recognition of these same emotions (Ekman et al., 1987).

Not everyone agrees that the results are strong enough to support the claim that basic emotions are "universally" recognized in the face (Russell, 1994). In general, however, from one end of the world to the other, a smile is a smile and a frown is a frown, and just about everyone knows what they mean—even when the expressions are "put on" by actors and not genuinely felt (Gosselin et al., 1995). Hillary Elfenbein and Nalini Ambady (2002) meta-analyzed ninety-seven emotion recognition studies
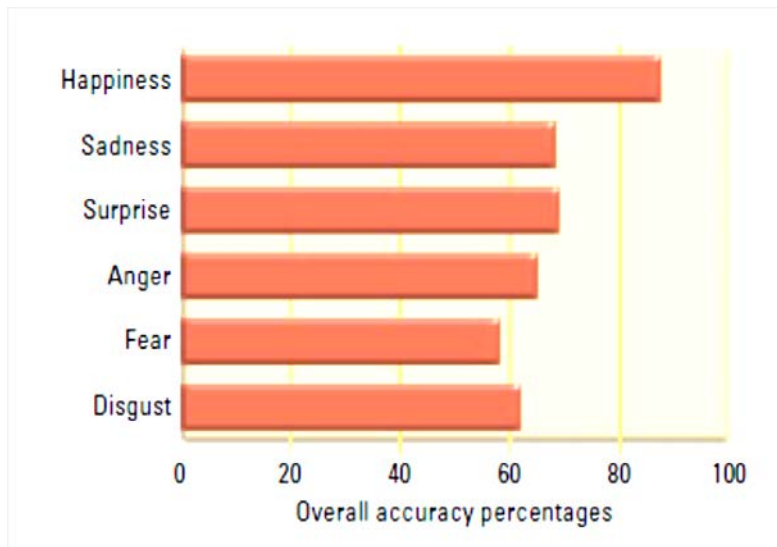
**FIGURE 10.12 How Good are People at Identifying Emotions in the Face?**
A meta-analysis of emotion recognition studies involving 22,148 participants from forty-two countries confirmed that people all over the world can recognize the six basic emotions from posed facial expressions (Elfenbein & Ambady, 2002).

involving a total of 22,148 people from forty-two different countries. As shown in Figure 10.12, they confirmed the main result that people can generally identify certain basic emotions from facial expressions. By comparing performances across different studies, they also discovered that people are 9 percent more accurate when they judge members of their own national, ethnic, or regional groups than when they judge members of other less familiar groups. In other words, we enjoy an "in-group advantage" when it comes to knowing how those who are closest to us are feeling. We also react to some emotional displays more than others. From an evolutionary standpoint, it is more adaptive to beware of someone who is angry, and likely to lash out in violence, than someone who is happy, a nonthreatening emotion. Studies have shown that angry faces arouse us and cause us to frown even when presented subliminally—without our awareness (Dimberg & Ohman, 1996; Dimberg et al., 2000).

Emotion is accompanied by changes in facial expression—even when these changes are subtle and cannot be seen with the naked eye. The human face has eighty muscles that can create more than seven thousand different expressions. To measure the spontaneous activity of these muscles and their links to emotion, many researchers use a physiological device known as the **facial electromyograph (EMG).** In facial EMG studies, participants are shown images that evoke positive or negative emotions, while electrodes attached to the face record the activity of various muscles (see Figure 10.13 on page 402). This research shows that images that elicit positive emotions such as joy, interest, and attraction increase activity in the cheek muscles; those that arouse negative emotions such as anger, distress, and fear spark activity in the forehead and brow area. Evidently, the muscles in the human face reveal smiles, frowns, and other expressions that are otherwise hidden from view (Dimberg, 1990; Tassinary & Cacioppo, 1992).

SENSORY FEEDBACK

Draw the corners of your mouth back and up and wrinkle your eye muscles. Relax. Now raise your eyebrows, open your eyes wide, and let your mouth drop open slightly. Relax. Now pull your brows down and together and clench your teeth. Relax. If you followed each of these directions, you would have appeared to others to be feeling first happy, then fearful, and finally angry. The question is, do these expressions affect how you actually feel?
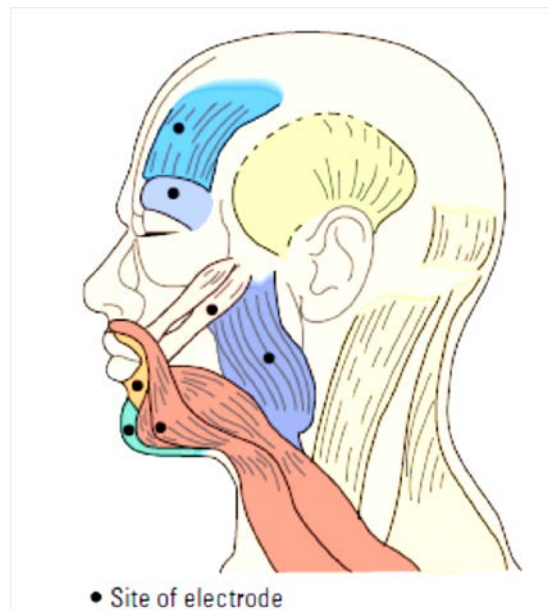
FIGURE 10.13 The Facial Electromyograph (EMG)
Electrodes placed on the face record activity in various muscles. These recordings reveal that positive emotions increase activity in the cheek muscles, and negative emotions increase activity in the forehead and brow areas (Cacioppo & Petty, 1981).

According to the **facial-feedback hypothesis,** an expression does more than simply reflect one's emotion—it actually triggers an emotional experience. In an interesting first test of this hypothesis, James Laird (1974) told college students that they would take part in an experiment on the activity of the facial muscles. After attaching electrodes to the face, he showed them a series of cartoons and asked them before each one to contract certain facial muscles in ways that made them smile or frown. The result: The students thought the material was funnier and reported feeling happier when they wore a smile than a frown. Similarly, other posed-expression studies show that people can also be induced to experience fear, anger, sadness, and disgust (Duclos et al., 1989). Together, this research suggests that facial expressions—though not *necessary* for the experience of emotion—can evoke and magnify certain emotional states (McIntosh, 1996).

Why does this occur? Laird believes that facial expressions activate emotion through a process of self-perception: "If I'm smiling, I must be happy." To test this hypothesis, Chris Kleinke and his colleagues (1998) asked people to emulate either the happy or angry facial expressions that were depicted in a series of pictures. Half the participants saw themselves in a mirror during the task; the others did not. Did these manipulations affect mood states? Yes. Compared to participants in a no-expression control group, those who put on happy faces felt better—and those who put on angry faces felt worse. As Laird would predict, these effects were most pronounced among participants who saw themselves in a mirror.

Other researchers speculate that there is a second possible reason for this effect, that perhaps expressions trigger an emotional experience by causing physiological changes in the brain (Izard, 1990). For example, Robert Zajonc (1993) proposed that smiling causes facial muscles to increase the flow of air-cooled blood to the brain, which has a pleasant effect by lowering the brain's temperature. Conversely, frowning decreases blood flow, which produces an unpleasant state by raising brain temperature. To demonstrate this mechanism, Zajonc and his colleagues (1989) asked participants to repeat certain vowels twenty times each, including the sounds *ah, e, u,* and the German vowel *ü.* As they uttered the sounds, temperature in the forehead was measured and participants reported on how they felt. The result: *ah* and *e* (vowel sounds that cause speakers to mimic smiling) lowered forehead temperature and elevated mood, whereas *u* and *ü* (vowels that cause speakers to mimic frowning) raised temperature and dampened mood. In other words, movement of the facial muscles influenced emotion even though participants didn't realize that they were wearing an expression. The lesson: If you want an emotional lift, just put on a happy face.

110

©VISION IAS
www.visionias.wordpress.com
Contact: +91-9650617807, +91-9968029039

©Numerons
www.numerons.com
Contact: +91-9560344245

VISION IAS

www.numerons.com