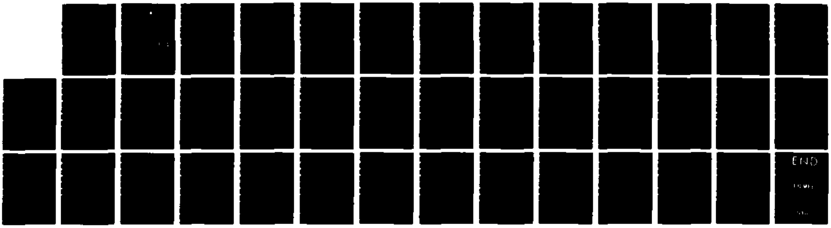
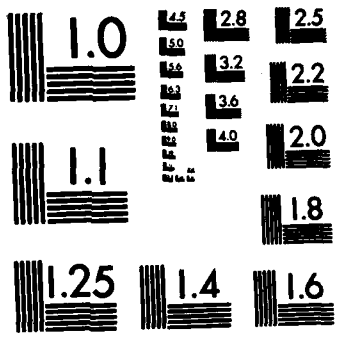
AD-A159 452 GUIDELINES FOR THE EVALUATION OF VISUAL APPROACH SLOPE INDICATORS(U) AERONAUTICAL RESEARCH LABS MELBOURNE (AUSTRALIA) J MILLAR SEP 84 ARL/SYS-R-32 1/1

UNCLASSIFIED F/G 1/4 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(2)

ARL-SYS-R-32 AR-003-962

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

AERONAUTICAL RESEARCH LABORATORIES

MELBOURNE, VICTORIA

SYSTEMS REPORT 32

GUIDELINES FOR THE EVALUATION OF VISUAL APPROACH SLOPE INDICATORS

by

JANE MILLAR

Approved for Public Release

COPY No

SEPTEMBER 1984

AR-003-962

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
AERONAUTICAL RESEARCH LABORATORIES

SYSTEMS REPORT 32

GUIDELINES FOR THE EVALUATION OF VISUAL APPROACH SLOPE INDICATORS

by

JANE MILLAR

SUMMARY

This report discusses techniques for evaluating the efficacy of Visual Approach Slope Indicators (VASIs) intended for transport aircraft landing at major runways. The discussion is predicated on the assumption that VASI signals should act as a substitute for the information the pilot sees in the natural scene, and a list of design requirements is developed accordingly. Several testing techniques and the utility of information likely to be obtained from each one are outlined. Various aspects of experimental design including the choice of aircraft, pilots, treatments and statistical analysis which can influence the validity of conclusions are highlighted. In many published evaluations to date there has been undue reliance on the subjective opinions of pilots instead of measurements describing flight paths in objective terms. Similarly, operational evaluations have been preferred to more informative experimental tests. Adequate testing of current VASIs could enable substantial improvements to be incorporated into VASI designs of the future, but unfortunately, previous tests have not always been adequate, restricting the pertinence of the findings and limiting the generality of conclusions which might otherwise have influenced subsequent designs.

POSTAL ADDRESS: Director, Aeronautical Research Laboratories,
Box 4331, P.O., Melbourne, Victoria, 3001, Australia

CONTENTS

1. INTRODUCTION

Visual Approach Slope Indicators (VASIs) are systems of lights which are installed alongside runways to signal to pilots visual information about the approach path. Since the first VASIs were devised, almost forty years ago, many different kinds have been developed to assist pilots landing a variety of aircraft under differing circumstances. Traditionally, attention has been directed mainly towards guiding large transport aircraft landing at major airports, although some consideration has been given to catering for general aviation operating at secondary airports, sea-going carrier landings and night military operations at makeshift helicopter pads and STOL runways.

Currently, two aids, T-VASIS and [Red-White] VASIS, are approved for turbojet (or similar) aircraft operations by the International Civil Aviation Organization (ICAO). The original designs of both these VASIs are still being used even though they date back further than twenty years. However, the expansion of aviation with the concurrent increasing diversity of new aircraft types has prompted numerous (at times up to ten) versions of Red-White VASIS being included in the ICAO standards (Annex 14) and one (tentative) modification to T-VASIS (described by Gregson 1978). Several of these variants are intended to suit secondary airports where pilots of smaller turbojet or STOL aircraft may benefit from some VASI guidance.

From time to time, different devices have also been suggested and recently, newer VASIs such as PAPI (Precision Approach Path Indicator) developed by Smith and Johnson at the Royal Aircraft Establishment and PLASI (Pulse Light Approach Slope Indicator) manufactured by the DeVore Corporation, have been marketed, partly with secondary airport users in mind. PAPI was initially developed for tactical military purposes, particularly for helicopters and STOL aircraft (CAA 1978), but subsequently it was proposed as being suitable for aircraft operating at major airports (Smith and Johnson 1976), and later presented to ICAO as a replacement for Red-White VASIS (Brown 1978) which has unsatisfactory guidance characteristics. Other possible uses of PAPI, after modification, include shipboard landings (Dawson 1979) and operations at secondary airports (e.g. Hald 1980). Similarly, PLASI has been suggested as useful for aircraft at major runways, secondary airports and oil rigs (De Vore 1981, 1982).

The continued development of different VASIs for the same applications, whether for the more traditional or the newer usage, is a matter of concern from the human factors viewpoint. If the older VASIs exhibit undesirable characteristics during operation or cannot satisfactorily be adapted to meet modern requirements, then it is reasonable to replace them by more appropriate devices. However, to ensure adequate performance and to avoid possible pilot confusion, it would be preferable if only one VASI (or perhaps two) was selected for each generic application. In the interests of standardisation, then, there is a need to choose between contenders.

The methods for choosing between VASIs for international-class operations of fixed-wing aircraft constitutes the topic of this report with the major emphasis being placed upon proper testing of the human-factor related qualities (i.e. the interplay between the pilot, the aircraft and the signalled information). Before proceeding with a description of techniques for evaluation, research about why pilots sometimes misjudge their approach is examined so as to derive a design philosophy for VASIs. Operational requirements which constrain VASI design, worthwhile ergonomic principles and hardware features are then listed to illustrate the numerous aspects requiring evaluation. Testing methods are outlined under four major headings here—two are devoted to laboratory studies, including the use of manned flight simulations, and two to field investigations. More frequently evaluations are conducted in the field. Therefore, considerable attention has been devoted below to difficulties which may arise when designing experiments or undertaking operational trials, assessing the reliability of the results and drawing conclusions. These same evaluative techniques and principles can be applied to VASIs for other applications, including specific military purposes.

2. DESIGN ASPECTS OF VASIs FOR INTERNATIONAL OPERATIONS

During an approach the pilot generally relies heavily on visual cues from the external scene to judge the flight path. However, pilots do not always perceive their glidepath as it truly is, and errors in visual judgement when tracking the glideslope are frequently responsible for, or precipitate, landing accidents (Lane and Cumming 1956, 1959). It is precisely these accidents that VASIs have been designed to prevent. Since VASIs rely on pilots interpreting visual information, as does a normal unaided approach to land, it is important to understand precisely just when, how, and even why, misinterpretations of three-dimensional relationships in space arise. This information can then be used to derive a design philosophy. Further, an understanding of the visual processes used when landing could ensure that similar undesirable perceptual features are not incorporated unwittingly into VASI signals.

2.1 Visual Perception During Landing

A good deal of research effort, particularly during those years when the first VASIs were designed, has been devoted towards trying to tease out the visual factors which are important to a pilot when landing. There are several avenues available for doing so. Of these, three were preferred in the earlier years before today's improvements in external displays for manned flight simulators—one way was to ask pilots about the conditions they found difficult, and which might or might not have resulted in reports of the circumstances reaching avaition authorities, another to investigate (both in the laboratory and in the field) the visual strategies taught during training and those preferred by experienced pilots, and the third, to undertake rather more academic experiments directed at fundamental aspects of visual perception and visual system organization.

Anecdotal pilot accounts and formal surveys of pilot opinion have revealed a variety of circumstances which can create perceptual difficulties. The list is extensive and encompasses topographical features of the surrounding terrain and the runways themselves, atmospheric visibility conditions and the view through aircraft transparencies. Specific factors implicated in misjudgements of approach paths include—a false horizon or a lack of a visible horizon; conditions where depth cues like interposition, texture or shading are absent or misleading (e.g. as in approaches over water or at country strips where there are no buildings nearby); an unusual ratio of width to length of the runway; a runway with a dip or a mound; a sloping runway; isolated or elevated lights on the ground at night; atmospheric haze, mist, rain, fog or a low cloud base; scratched or degraded windscreens; optically distorting windscreens or a restricted field of view from the pilot's seat; and, numerous other conditions. Accordingly, ICAO has recognized that pilots may experience perceptual difficulties where the available topographical cues are minimal or misleading and has recommended that a VASI be installed where these conditions exist (Annex 14, Section 5.3.6.1(b)).

Experiments incorporating displays of airport scenes at night have also produced findings which suggest that pilots have difficulty in judging suitably safe and accurate approach paths from their view of the simulated runway and surrounds. Poor judgement of a simulated approach is especially likely if mountainous terrain is depicted by elevated lights, and/or the foreground is sparsely lit, or if the runway slopes, and/or if the runway shape (e.g. width to length ratio) is unfamiliar to the pilot (Kraft and Elworth 1968; Kraft 1969; Stout and Stephens 1975; Mertens 1978, 1981; Lewis and Mertens 1979; Mertens and Lewis 1982). Several other simulation studies, some not specifically directed at these issues, have also reported similar findings of pilot misjudgements under a variety of experimental conditions and depicted scenes.

Many of the conditions outlined above are often said to cause the rather loosely termed phenomena of "visual illusions" (see for instance Vinache 1947; Pitts 1967; Armstrong, Bertone and Kahn 1972; Iwataki 1973; Crymble 1977; Sams 1980). (Note that certain psychologists may also refer to illusions as "nonveridical" perception.) Categorising erroneous perception as illusory, however, does not provide us with an explanation of the reasons why, but merely broadly describes a multitude of phenomena which may or may not be caused by the same factor(s). For example, there are almost as many explanations for why geometric illusions occur as there are illusions (see general psychology texts such as Murch 1973). Nevertheless, recent research has suggested that there could be a common factor underlying misjudgements of depths as varied as those that occur when viewing the moon at night, or when viewing simulated displays or when landing aircraft.

Roscoe and his colleagues have, in a series of experiments designed to investigate the perception of optical and computer displays, recorded the accommodative responses (i.e. focusing position) of individual's eyes to visual scenes. Consistently, the findings have indicated that inappropriate accommodation with collimated images or remote real objects is associated with inaccurate estimates of absolute distances. Roscoe has suggested that similar accommodative responses cause pilots to misjudge distances in real aircraft and so perhaps generate short landings (Roscoe 1979, 1980, 1981). Should this prove to be true, then Roscoe suggests that it may be possible to improve distance perception by correcting accommodation with spectacles. Several authors (Kraft, Hennessy and others) have previously remarked that glasses may be beneficial when flying (Roscoe 1980).

Roscoe's hypothesis is an extremely attractive one, but it remains tentative at this stage until further evidence accumulates. His solution also requires further testing, particularly because there may be degradation of other visual system functions as a consequence of wearing corrective lenses for distance viewing. For instance, vision of nearer objects may be disrupted. Roscoe comments elsewhere in his 1981 paper that the human visual system is remarkably tolerant of blurred images. Hence optical correction for viewing distant scenes may cause pilots to unconsciously accept fuzzy images of closer objects (e.g. instruments) which are difficult to resolve readily when the pupil provides a shallow depth of field (in dim illumination) without further accommodative effort. Perhaps bi- or even multi-focal lenses could be useful, but these lenses contribute their own perceptual complications as readers who wear them will remember from their first efforts to acclimatize.

Despite intensive exploration, the visual processes important in the landing phase are still not very well understood. Even though it is thought that most strategy-related cues (such as the H-distance or the method of aligning objects in certain portions of the windscreen) have been broadly categorized (see Lane and Cumming 1956; Mertens 1981), neither their hierarchy of importance nor the interactions between them can be confidently stated. Experimental investigations have so far not supported pilot opinion, nor indeed theoretical predictions, about the importance or supremacy of specific cues when the visual scenes are complex, although there have been several investigations which have manipulated certain variables related to one or two cues (for instance the runway shape) while maintaining others in simple scenes (see Mertens (1981) for example). Further, the classical academic psychology method of reducing the visual scene to simple, abstract perceptual components (for example monocular cues of shading, occlusion, perspective, etc. and binocular cues such as convergence and disparity) has also not proved very fruitful in understanding the pilot's task during landing. Although many such building block cues are recognized (see for instance Winterberg, Brictson and Wulfeck 1966; Riordan 1974; Regan 1980), and a few attempts have been made to incorporate these cues into models of the pilot's task during landing (e.g. Wewerinke 1978), the analyses are so far rudimentary. The individual relevance and interactions between these cues have been difficult to relate pragmatically; this is a fact that Regan (1980) attributes to the traditional introspective "bottoms-up" analytical method which mainly considers static scenes rather than the dynamic.

One theoretical analysis uses the "tops-down" method (currently a method being advocated by Regan, amongst others, which starts with the entire scene instead of its components) and considers the optical flow patterns induced during dynamic situations. The general concept was called motion perspective by Gibson (see Gibson 1950). Gibson nominated the focus of expansion as the most important motion perspective cue (Gibson, Olum and Rosenblatt 1955). This idea has received further attention by others (see for instance Hochberg and Smith 1955) and is frequently considered to be a satisfactory explanation of the pilot's visual strategy. However, laboratory experiments by Johnston, White and Cumming (1973) have not confirmed Gibson's proposition. These researchers concluded that the focus of expansion is a relatively weak cue for tracking the glideslope because subjects found it difficult to distinguish accurately in the laboratory. Winterberg *et al.* (1966), on the basis of a mathematical analysis, also thought that the cue would be of little use to pilots approaching towards a carrier at night. However, recently interest has been revived in determining the usefulness of the various optical flow patterns induced during approach and landing because of requirements to design adequate simulations (Owens and Jensen 1981). Further interest in flow patterns has been generated by Regan and his colleagues (particularly Beverley) who have found that the binocular channels in the visual system are surprisingly sensitive to movement (see Regan 1980). Currently, though, because perception is an extremely complex process, it has so far defied theoretical attempts at global synthesis, and the visual processes by which pilots land aircraft remain basically unknown.

2.2 The Major Premise Underlying VASI Design

Considering that we do not have a good grasp of the visual processes involved, it is therefore not surprising that VASIs have usually not been designed specifically to ENHANCE the visual world by emphasizing the most important cues to a pilot approaching to land. There are, however, other types of visual approach aids which do attempt to capitalize on perceptual abilities derived from those which seem to be used during landing. For instance, it has often been suggested that one way pilots can judge their approach angle is from the appearance of the runway. This cue has been emphasized in some configurations of pre-threshold approach lighting and in the idea of painting diamonds or other figures on the tarmac to give extra geometry for judging the approach angle. Opinions are divided about the utility of diamonds for pilots at secondary airports, but extrapolating from the results of Swaroop and Ashworth (1978) it is clear that sufficient approach accuracy could not be achieved by pilots of international-class aircraft. These are the only examples of approach aids which rely primarily on enhancing "natural" cues. VASI information, on the other hand, is usually provided in the form of a symbolic display, similar to some of those shown by aircraft instruments.

It appears then, that instead of enhancing important cues, VASI guidance has been obtained using a different design philosophy. In essence, many VASIs have been intended to REPLACE the visual scene, by containing sufficient information from which a pilot may accurately judge the aircraft's glidepath.

Undoubtedly, some scientists would argue that, in general, the intention underlying VASI design is only to SUPPLEMENT the visual world by providing extra guidance to which pilots may refer if warranted. However, it is evident from the many sources of information that we have about pilot perception during approach and landing phases that judgements of a proper path are sometimes highly unreliable and therefore it is unrealistic to expect that a pilot could always recognize situations where he might require assistance from a VASI. Further, the research of Lewis and Mertens (1979) suggests that pilots may not make observations of certain VASIs frequently enough to know whether their aircraft is remaining within acceptable limits.

Accordingly, the major premise underlying the design of each VASI is an important consideration. Further aspects of design are considered below; these are based on the major premise that a VASI should be designed as a replacement, instead of a supplement, for the information which pilots otherwise obtain (albeit sometimes imprecisely) from the real world. These lists have been compiled to illustrate that design requirements for VASIs are already quite comprehensive and that each item requires assessment during some stage of an evaluation programme.

2.3 Design Requirements for VASIs

The VASI must also fulfil requirements other than those based purely upon the psychological capabilities of the pilot. These were categorized by Whittenburg, Vaughan, Havron and Cavonius (1964) as:

(1) operational requirements,

and

(2) imposed constraints (e.g. geographic, economic and meteorological).

These two categoeies do not distinguish the features of VASIs based on psychological attributes of the pilot, and so different categories have been used below in order to stress the necessity of providing the pilot with clear and accurate information. Economic factors are not discussed; the trade-offs involved in deciding whether to install a VASI at any particular airport were reviewed for the USA context by Roman (1977), to whom interested readers are referred.

The following lists have been compiled from the suggestions of several authors (e.g. Lane and Cumming 1956, 1959; Sparke 1958; Morrall 1960; Baxter, Day and Lane 1960; Baxter, Cumming, Day and Lane 1960; Cumming 1962; Smith and Johnson 1976), although the liberty has been taken of rewording, reordering, up-dating and adding to these suggestions. Many of the statements are unquantified because there is still insufficient knowledge to assign values. The order in which the various requirements are listed does not imply merit priority; to some extent value judgements need to be imposed by the evaluator when assessing the relative significance of each factor.

2.3.1 Information content

A pilot approaching to land requires certain information from the visual world to judge whether the aircraft trajectory will reach the aiming point on the runway. The current consensus seems to denote vertical information as the most valuable feature of VASI guidance. It is usually assumed that lateral alignment information can be derived from the runway shape (including texture or lighting patterns), and although the guidance precision is unquantified, this is considered to be adequate judging from the low incidence of accidents involving these misjudgements. However, theoretically, ambiguities may arise if lateral cues are not present in the VASI signals because pilots may confuse aircraft roll with deviations from centreline when the horizon is not visible (Cumming and Lane 1957).

The information conveyed unequivocally by a VASI for international turbojet aircraft should:

(i) indicate the vertical position of the aircraft in relation to the nominal glideslope;

(ii) provide information about the rate of change in position;

(iii) give a clear indication about the direction of any change;

(iv) indicate deviation in azimuth between the aircraft's position and the extended runway centreline;

and

(v) provide adequate roll guidance information.

2.3.2 Operational requirements

Most of the following operational requirements have been extracted from ICAO Annex 14 which deals with VASIs for international aircraft.

(i) The information should be provided over a useful range:

(*a*) at least as far as the ICAO-approved minimum of 7·4 km by day and by night in clear weather;

and

(*b*) down to a suitably low height above the runway.

(ii) The information provided should be in a form sufficiently adaptable to allow:

(*a*) wheel clearance over the threshold in accordance with ICAO standards;

and

(*b*) non-standard approach paths to be followed safely and accurately when necessary.

and

(iii) The effect of adverse meteorological conditions on the integrity and visibility of the signal information should be minimal.

2.3.3 Ergonomic principles

It is not only necessary to provide glidepath information to the pilot, but also to ensure that the information is in a form that can be comprehended readily and reacted to effectively. It is evident from several sources of knowledge we have about cognitive processes that sometimes people can make wrong decisions by unconsciously ignoring pertinent information, especially when they are too busy, confused (e.g. disorientated), distracted or fatigued. If pilots perceive conflicting sources of visual information in these situations, it is possible that they may accept their perception of the external world (which may be incorrect) instead of the VASI signals.

Therefore, VASI information should be displayed so that the system is:

(i) failsafe;

(ii) reliable;

(iii) easy to learn;

(iv) easy to interpret;

(v) unambiguous;

(vi) distinct from any other VASI;

and

(vii) easily identifiable from other airport lighting (without producing too much glare).

The information should also be displayed in a manner which assists the pilot to guide the aircraft accurately down the nominal glidepath. Psychological research extending over many years has revealed that in order to display the VASI information to the pilot in an ergonomically sound manner, there are certain features which should be incorporated. Current knowledge about the strengths and weaknesses in several VASI designs (e.g. Precision Visual Glidepath and the conceptually related "meatball" systems which are usually installed on carriers; Red-White VASIS; T-VASIS and its precursor, TEE; and the tri-sectored VASIs like AAI, Bardic, GAIL, etc.) suggests that the following features are highly desirable.

To assist the pilot when following the indicated glideslope the signals should have:

(i) a sufficient number of guidance categories to provide adequate (precise) feedback;

(ii) a "sensitivity" compatible with all the possible pilot-aircraft combinations likely to use the aid;

and

(iii) compatibility of the signals with other sources of glideslope information such as the ILS signal.

Quick and reliable response from the pilot to changes in the VASI signal should be facilitated by the use of:

(i) a "director-type" code;

(ii) an "all is well" signal, augmented by extra cues only when deviations occur;

(iii) an unmistakable warning signal for an unacceptably or dangerously low approach—preferably distinct from the normal signal;

and

(iv) pattern perception rather than colour coding or other specific visual abilities (e.g. vernier acuity).

2.3.4 Aspects of the hardware

The equipment that generates the signals should also be:

(i) failsafe;

(ii) reliable;

(iii) a minimum obstruction hazard;

(iv) robust and able to withstand jet blast;

(v) capable of variable output in light intensity without degrading the signal discriminability;

and

(vi) unaffected (or minimally so) by interference from natural causes, e.g. build-up of snow, intense solar radiation load, accumulation of dust, etc.

2.3.5 Additional constraints

Recently, Smith and Johnson (1976) suggested that a VASI design should encompass certain further constraints, viz.:

(i) The system should be so designed that it can be installed without the aid of sophisticated site survey equipment and should not require a flight check before use.

(ii) The equipment should be capable of being left unattended for long periods.

(iii) Unit and systems costs should not be significantly greater than the costs for current in-service systems.

and

(iv) The system should have the capability of adequately supporting all current types of operation and of meeting any future operational needs such as steep approaches.

These four points are mentioned separately because, according to a recent paper by Smith (1981), these constraints were specifically applied during the design phase of PAPI. As far as the current writer is aware, these ideas are new and therefore it might be worthwhile commenting here on their general applicability to commercial fixed-wing operations.

The first two, (i) and (ii), are likely to be military requirements and probably have more application to secondary rather than major airports. However, there seems to be little reason why an adequate ground checking system could not be developed for all the current aids, as apparently it has for PAPI, but the requirement not to flight check before initial use seems to be unnecessarily restrictive. Furthermore, it is sensible to ensure that VASI equipment does not require excessive maintenance even at a major airport where staff are presumably available to undertake the work. The third point, (iii), is pragmatic in these days of stringent financial restraints, but the relative costs need to be weighed against the efficacy of the VASIs in the comparison. The price of VASI units is minute in comparison to savings from the prevention of potential accidents. Not everyone holds this opinion however, and so the reader is referred to Smith and Johnson (1976) or to the British presentations to ICAO (see Millar (1984) for a full bibliography).

The fourth point, (iv), entails a rather more ambitious aim and as shown in the ensuing sections of this paper, an aim which is unlikely to be fulfilled without further test and evaluation involving the new conditions.

3. METHODS FOR TESTING AND EVALUATING VASIs

There are several techniques available for gauging the usefulness of VASI guidance to pilots and for comparing different VASIs. The data come from three main sources; the laboratory, where equipment ranges in complexity from pencil and paper, slide projectors, etc. to sophisticated manned flight simulators; field trials where aircraft are flown under either experimental or operational conditions; and the routine operational environment which provides incident and accident reports. In the following pages these techniques are reviewed, placing emphasis on the advantages and disadvantages found with each. Experiments in the laboratory are considered in two sections below, with the traditional, simple apparatus treated separately from manned flight simulators because, even though advantages are shared, substantially different ideas about their relative efficiency and about the conclusions which can be derived from each are commonly held. Background information has been drawn from the comments and suggestions in articles reviewing the evaluation literature and from research reports about VASI guidance for large fixed-wing aircraft at major airports (e.g. Lane and Cumming 1956; Baxter, Cumming, Day and Lane 1960; Morrall 1960; Cumming 1962; Vaughan, Luce and Kassebaum 1962; Watters, Rollins, Frey and Cavonius 1964; Whittenburg *et al.* 1964; Winterberg, Brictson and Wulfeck 1966; and H. J. Clark 1968; amongst others).

3.1 Laboratory Studies

The laboratory is an obvious place to begin any evaluation. Apart from testing equipment and hardware in the laboratory, the controlled environment is suited to investigating human factor features of a VASI design. The ability to control the experimental environment (by keeping chosen conditions constant while manipulating others) is a major advantage because unwanted influences which may occur unpredictably in the real world can easily confuse (confound) or mask otherwise unnoticeable, but important, psychological features. Aspects such as reaction time, visual discrimination, decision making and tracking performance which are required in response to VASI signals may be tested either alone or in combination with one another to assess whether they are appropriately matched to human sensory, perceptual, motor and cognitive skills. Usually, the experimental apparatus is relatively simple (e.g. tachistoscopes (viewing boxes), 35 mm projectors and line drawings are often used for displaying stimuli) by comparison with the complexity of modern aircraft systems, and the experiments themselves can be highly conceptualized representations of flying and associated tasks.

A particular example of a successful application of laboratory results to VASI evaluation was based upon the results of an experiment by Hunt (1961) who found that better tracking performance can be achieved when more than three categories of information are provided to the tracker. Hunt's task required the subjects to track the movement of a spot (generated by an analogue computer) which was seen in various positions through a slot. Cumming (1962) used these results to predict that the imprecise guidance of Red-White VASIS (with three categories) would be even less appropriate when newer aircraft (presumably jet-engined) with slower response times were introduced. His prediction was subsequently vindicated by the concern which has been expressed about Red-White VASIS at ICAO and the modern flight trial reports by Smith and Johnson (1973, 1976) which echoed similar results to those found in the earlier Australian work.

In addition, laboratory experiments are useful for developing general design principles. The following two examples illustrate this point. Whittenburg *et al.* (1964) reviewed some of the laboratory experiments they and their colleagues (e.g. Vaughan, Rollins and Luce 1963) undertook when investigating lighting suitable for helicopter approaches, and they derived a list of potential rules for spacings between lights etc. at airfields; the designers of T-VASIS were influenced by Hunt's tracking accuracy results and so incorporated seven primary signals into its code (Cumming 1962).

Studies in the laboratory can also contribute to an understanding of events that may go unnoticed in normal aviation operations but which may be important when a particularly high workload is imposed. The classic way to produce a high workload in the laboratory is to introduce another task, called a secondary (or side) task, and to measure changes in performance on the primary task. However, there is some concern that adding artificial or stylized tasks (to what may already be a fairly unrealistic situation) can cause artefactual interactions between the tasks (Chapanis 1967). Therefore the interpretation of the data and the conclusions drawn from such an experiment may be misleading. These interactions can be avoided by eliminating the secondary task and making the primary one harder, as the following example demonstrates. The task required the subject to remain within the "on-glideslope" corridor of a VASI as accurately as possible. The pilot subject was told that he was controlling a large aircraft, the dynamics of which mediated the time between his response and changes in the display of the VASI code on a monitor. All visual cues found in flying (such as movement of the lights in the screen representing changes in the flight path as viewed through the windscreen) were deliberately suppressed, and therefore intense concentration was required from the subjects (see Millar and Selway 1981 for other details of the experimental design). When the performance on two similar VASIs was compared to another, conceptually different VASI, it was observed that control reversals (i.e. the subject pulled back on the stick instead of pushing forward or vice versa as was appropriate) were more common with the latter type. Although further work would be needed to establish the relevance of these observations, an area of potential difficulty was delineated in this simple experiment.

Other simple simulations are also possible. The experiments of Baxter, Day and Lane (1960) (see also Day, Baxter and Lane 1960) who investigated the visual acuity required when judging the Precision Visual Glidepath (PVG) signals, illustrate an effective simulation. Thresholds for detecting misalignments of the PVG bars, and those which professional pilots considered substantial enough to correct in flight, were obtained when the subjects viewed model boards. The display was static and showed the PVG at various ranges and approach angles in good and poor visibility. (For instance, rain was simulated by spraying water over the viewing window.) In further experiments it was demonstrated that the laboratory model was a reasonable representation of the PVG and that the laboratory conditions produced similar perceptual degradation to vision as does viewing the PVG when there is reduced visibility in the field. The authors then explored further perceptual effects (such as the influence of adding runway lighting to the display) in the laboratory and predicted whether pilots could obtain usable information under similar field conditions.

Despite these possibilities, previously the laboratory has often been overlooked or needlessly underrated. Some of the reluctance to capitalize on the advantages may relate to a degree of uncertainty about the extrapolation of the results to the "real" world. This is not altogether difficult to understand as some experiments seem to have tenuous links with flying and especially when errors of extrapolation have occurred in the past. Chapanis (1967) mentions several unsuccessful attempts where conclusions based on laboratory results have been incorrectly applied to real world tasks of various kinds. Another notable example concerns the choice of colour for instruments and floodlights used to illuminate the cockpit during night flying and this example is illustrated below.

Laboratory studies on the visual system have indicated that the rod receptors in the retina, which mainly mediate vision at night when light levels are low (i.e. the scotopic range of light levels), are relatively insensitive to wavelengths lying in the red end of the spectrum. Once rods are exposed to higher intensity light they become "bleached" (literally) and take some time to readapt to lower light levels, depending on the degree of exposure to light. In comparison, the cone receptors, which are mainly activated in the day when luminance levels are higher (the photopic range), are necessary for seeing the fine detail such as that on instrument dials. Cones are sensitive to red wavelengths. Therefore for night flying it seemed that it would be advantageous to provide the cones with higher intensity red light illuminating the cockpit so that the instruments were visible, leaving the rods in their dark-adapted state for vision outside the cockpit.

Red cockpit lighting, however, has not been successful as one may have expected. White lighting has often proved superior in a variety of night flying operations, principally because red light has a number of undesirable features including the difficulty pilots have when reading colour-coded details (and interpreting them correctly) on maps and instruments and the tendency for some people to feel uneasy or disorientated in red surrounds (Johnson and Poston 1976). Further, the advantages of red light are lost in the slightly brighter conditions (either the mesopic—twilight—or the low end of the photopic range) experienced frequently during many operational flights (AGARD 1967).

This example illustrates the complications which may arise if insufficient attention is paid to the possible interactions between competing demands of the real world tasks. To avoid missing important connections it might seem better to conduct the experiments in the field, but this environment contains many variables (e.g. weather, other air traffic, radio calls etc.) which cannot be controlled by the experimenter. Uncontrolled variables differentially influence the performance parameters, resulting in increased variances which coarsen the measures and so may swamp indicators of difficulties for the pilot. These difficulties may normally be innocuous because the pilot can compensate for deficiencies by adapting his behaviour (Watters *et al.* 1964), but they may become important when extra-ordinary effort is required (for instance in an emergency). However, situations such as these happen infrequently and might not occur during field trials. Without these major potentiating events, the subtle interactions could easily be masked by the variability of other influences on the performance measures and so be missed during a field evaluation. This is where laboratory experiments which parcel certain factors, holding some constant and varying others, can excel at differentiating between features of one or more VASI designs.

3.2 Aircraft Simulator Studies

The introduction to manned flight simulators of external displays which are visually compelling has provided another opportunity for studying VASIs in a controlled experimental environment and their use seems to be supplanting that of the traditional laboratory apparatus. Modern simulators have further advantages over less complex laboratory apparatus, allowing the experimenter to investigate at will the effects and interactions between VASI signalling systems and many other variables, especially dynamic ones, under quasi-operational conditions. Simulators are also excellent tools for safely (and cheaply) studying the effects of high workloads and critical or potentially dangerous situations akin to those encountered during flying (like severe wind structure or engine failure). Additionally, data collection may be facilitated by the automatic recording capabilities thus providing the scope for expediently investigating many different VASI configurations.

Considering all these attractive features, it is therefore tempting to evaluate VASIs solely through simulation, without running the more time-consuming and costly field trials. However, as with other laboratory-based experiments, again the issue of extrapolation to the real world is raised. The huge investments by the major airlines in simulators for training, checking and accrediting pilot time in lieu of substantial amounts of aircraft time attests to the general faith in simulation to depict flight adequately. Nevertheless, despite this apparent congruence, the results from simulator trials have "face-validity" (on first appearance) not "external validity" (i.e. extrapolation to the real world), enabling only the formulation of hypotheses, and not statements, about performance in the real world. Concern about extrapolation has been expressed on a number of occasions in the simulator literature, but seldom, if ever, in the VASI literature, so it is worthwhile briefly considering a few of the issues here.

The improvements in simulator technology to achieve more realism received their main impetus from requirements to find less expensive but adequate training tools to circumvent the necessity of teaching pilots in aircraft (Roscoe 1981). The needs and aims are therefore somewhat different in the two applications; in training paradigms (models) we are searching only for a situation that will enable a transfer of skills to flying an aircraft and not for a faithful representation of the real flight situations as we might require when testing VASIs.

The designers of simulator hardware hope to achieve perceptual effects with enough correspondence, or similarity (although not necessarily realism, at least in trainers), to those experienced during flight without, of course, actually flying. To do so, a knowledge of human perceptual and cognitive abilities and the relevant cues for the task(s) being trained is required so that adequate simulator performance is achieved without wasting computing resources on providing stimuli man can not sense or which are not useful for transfer of training effectiveness. Since our understanding of perceptual processes is still incomplete, approximations and limitations certainly have been incorporated into the models used for deriving simulator systems (AGARD 1980). And furthermore, psychological aspects which will prove to be vital as research continues, but which are so far undiscovered, may not be currently simulated. (An illustration of a recent discovery is that of the ability to detect movement through the binocular (stereoscopic) part of the human visual system (which was mentioned above—see Regan 1980).)

Although the trend is for more and more accurate representations of flight dynamics, motion, visuals, etc., it is not at all clear that sufficient realism has been achieved so far to permit a sufficiently exhaustive test of any VASI and thus avoid the necessity of conducting field trials. And in fact regardless of the standards achieved, even in the longer term, the results from simulators may never be a suitable substitution for field trials because, as a recent AGARD committee pointed out, absolute fidelity can only be obtained by actually duplicating the real world (AGARD 1981); this effectively rules out simulation as a substitute for the real world environment.

The authors of the 1981 AGARD paper mentioned the integration of motion bases as an illustration of their argument, underlying that motion fidelity can only truly be achieved by moving the simulator cabin through space in the same way an aircraft does; although it is not generally recognized, a similar argument applies to simulator visual systems. While issues about motion bases are debated hotly (e.g. Roscoe 1980; AGARD 1980), and motion may or may not be particularly realistic or even necessary depending upon why the simulator is being used, it is clear that a VASI evaluation requires at least an accurate visual representation of the aids under test and the (visual) perceptual conditions which could be experienced in flight.

Achieving an accurate simulated representation of visual conditions does not necessarily imply that the relevant visual conditions should be duplicated, however. The 1980 AGARD committee which investigated questions of fidelity for training purposes made this point when distinguishing between two classes of simulation qualities: OBJECTIVE and PERCEPTUAL fidelity. Objective fidelity refers to the accuracy of duplicating inanimate characteristics (including equipment, flight dynamics, aircraft behaviour, etc.), and the more accuracy achieved here, the greater confidence we can have when extrapolating to flight (i.e. the external validity). The second class, that of the perceptual fidelity, involves the degree to which the human can perceive the accuracy of the simulation in relation to its real live counterpart. These two classes are not as distinct as the AGARD definition implies because the performance of the pilot is dependent on how he senses the simulator environment including aircraft dynamics, etc. However for the purposes of our discussion we will treat them as separate issues and deal only with visual perception.

Chase (1971) showed that pilots' subjective evaluations of the fidelity of visual displays for simulators correlated in the main with their performance during "landing trials"; less accuracy was obtained on a number of parameters with those displays which the pilots rated as poor. While today's simulators may incorporate better visual representations than those studied by Chase, there are several aspects of visual system hardware which generate inadequate perceptual characteristics and therefore obviously limit their applications in VASI evaluations. For instance, the visual scenes displayed by computer generated imagery (CGI) usually limit the representations to dusk or night-time simulations because day-time requires an immense amount of detail to be incorporated. Day-time details needing computational power beyond current resources include gradients of textures which are considered to be of prime importance in the perception of depth and movement (Regan 1980) and which do improve accuracy of simulated landings if present (Buckland, Munroe and Mehrer 1980). Shadows and shading also are challenging to reproduce as is hiding the sides of surfaces that should not be seen in a 2-D projection of a 3-D object. Image quantisation (i.e. parcelling) may make edges of objects appear serrated instead of straight. Frequently objects are stylised or abstracted and appear cartoon-like or exhibit jerky movements. Depth relationships may also be unrealistic because of the difficulties mentioned above in reproducing monocular cues adequately, and because stereoscopic images are usually absent in displays using CGI in almost all simulators, except for an in-flight tanker display (AGARD 1981) and maybe others currently under development.

Night-time CGI simulations do not require such detailed images so many of these difficulties could be overcome, perhaps providing suitable conditions for evaluating VASIs, if it were not for several important limitations which often apply when depicting these scenes as well. For example, frequently simulator visual systems are restricted in their variety and saturation of colours, are unable to display the apparent fading of colours with simulated distance or are limited in their realistic portrayal of some atmospheric conditions (e.g. mist, rain, fog, haze, etc.). All of these are necessary conditions for properly testing colour-coded VASIs.

Simulations of Red-White VASIS, for instance, sometimes do not include a representation of the pink transition zone (e.g. Lewis and Mertens 1979) which is sometimes claimed to be a feature of the guidance afforded to pilots (see Gates 1970). Further inadequacies arise from displays that do not accurately model the filtering and scattering effects of common atmospheric aerosols (e.g. conditions of dust, smoke, smog or rain, etc.) because the effects on pilots' colour discrimination would not be properly represented. Yet it is known that atmospheric aerosols have a considerable influence on the colour perceived by the viewer (Clark and Gordon 1981) and it is likely that performance of pilots will deteriorate when the signalling code cannot be readily distinguished. The signals from Red-White VASIS are notoriously difficult to perceive in some sunny viewing conditions (Smith and Johnson 1973; 1976) and at a distance, but these are rarely represented in simulations. Many other VASIs (e.g. AAI, Bardic, PLASI and PAPI) also incorporate colour coding, and so presumably similar comments hold true for simulations of their codes too. An assessment of these VASIs would require day-time conditions to be simulated.

Perceptual anomalies may also arise in simulations. Various difficulties with raster cameras used to scan model boards can produce strange perceptual effects, during both simulated day and night, but often pilots do not notice them, and it is not known whether pilot behaviour is influenced by them (AGARD 1980). However, light trails perceived during image translation or from phosphor persistence may give some information about the direction of movement when they converge on the aiming point, providing an extra clue not observed in the real world. Various anomalies are also present in CGI images and one mentioned by an AGARD committee (AGARD 1980) could give very good, but unwanted, glideslope information in an experiment. This cue can be generated in CGI images and video views of model boards by distant features which occupy only a few picture elements (pixels), disappearing as the descent falls below a given relative angle and reappearing when the angle increases.

Another concern with using simulators to evaluate VASIs is the extent to which pilot behaviour corresponds with similar behaviour in an aircraft. The literature abounds with experiments concerning the efficacy of training in simulators (see Roscoe 1980) which show that transfer of skills usually occurs in a positive manner from simulator to aircraft, indicating that there is substantial correspondence between the two. However, for VASI assessments about acceptability, this correspondence may not be adequate.

Experiments directly exploring the correspondence between pilot behaviour in simulated and actual flight were not found by the writer's search of the VASI literature, but one unrelated example which explored drug interactions showed that simulator trials may not be quite accurate indicators of pilot performance during actual flying (Billings, Gerke and Wick 1975). These researchers found drug-induced performance decrements and learning effects in the simulator, but failed to find the drug effects to the same degree or consistently across all experimental parameters in the air. Further, variability in aircraft control was about 50 per cent greater during the flight trials. Although there were procedural inequalities between the two experiments, the authors considered that there was sufficient justification to speculate about why the results were different. They thought that possibly motivational factors reduced the action of the drug or assisted the pilot to compensate for any ill-effects during flight, reducing the main effects on flying skills. While they did not comment upon the other discrepancies, it is probable that the increased variability observed during the flight trials may have been caused by the loss of experimental control with unbalanced random events influencing the measures. (A statistically more powerful field experiment, which—in effect—strengthened the drug effects may have better replicated the laboratory results.) Further, the fact that the magnitudes of the main drug effects were different is inconsequential, the finding of main effects which were significant in both experiments is the important point. (Note that the statistical test used by the authors tells us about only the direction of any effect and not about its degree.)

Nevertheless, these results serve as a warning that certain pilot behaviour observed in simulators may not be evident during actual flight and, conversely, behaviour in an aircraft may not be duplicated in a simulator. According to Billings, Gerke and Wick (1975), "flying" a simulator requires different strategies to flying an aircraft, so that caution is warranted when extrapolating from simulator results to those effects which may be expected during flight. Another example is provided by the experiment of Fischer, Haines and Price (1980). They suggested that pilots did have difficulty ignoring the fact that approaches were being made every few minutes in a fixed-base device. The experimenters thought that the pilots may have been concentrating on the experimental head-up display (HUD) rather than on "landing safely", so over-emphasizing the assistance which might be obtained from the HUD in flight.

Motivational differences stemming from the relative consequences of making an error (in say navigation or landing) may also reflect in pilot performance measures. After all in a simulator the consequences of "crashing" are not nearly as severe as doing the same thing in an aircraft; only pride may be hurt in the first instance. As yet, the effects on performance of these motivational factors have not been extensively investigated, although phenomena such as "simulator complacency" have been recognised (e.g. Foushee 1981). Another psychological feature of pilot behaviour in simulators was mentioned by Fischer, Haines and Price (1980) who suggested that pilots form "expectancy sets" about encounters and hence may not expect runway obstacles in simulated scenes unlike in real aircraft where obstructions are more probable.

As we do not know what significance inaccuracies in the representation of flight, including visual factors, and motivational factors included in the task of "flying" have on the performance of pilots in simulators compared to that achieved in flight, it would be unwise to rely on the results of simulator evaluations to make categorical statements about the acceptability (or otherwise) of a particular VASI. Comparative judgements between VASI designs, on the other hand, may be facilitated by using the simulator prior to flight trials provided that there is reasonable congruency between the experimental situation and the real world, and that important interacting variables are accounted for. If used judiciously, in conjunction with developmental field trials, manned flight simulations have great potential for delineating specifications for future VASI designs.

3.3 Field Trials

It is possible to conduct two main types of formal evaluative trials in the field—either an experimental test or an operational test. In an experimental test the investigator deliberately selects a subset of pilots flying particular aircraft under a tight experimental regime. Operational tests, on the other hand, observe routine flying operations.

Generally there has been a marked preference for evaluating VASIs in the field under operational conditions (for instance PAPI—see Millar 1984). This preference, however, is not necessarily justified; often the "face validity" achieved is accepted as the basis for decisions (Cumming 1962), even though the events of interest occur infrequently and therefore large numbers of recorded trials are required before a trend can be predicted with any reliability. This sampling issue is seldom considered and often the results from only a few trials are collected.

The experimental field evaluation is by far the most powerful technique because irrelevant factors which may confound operational evaluations can be better controlled by experimental methodology (Cumming 1962). Therefore, in an experiment it should be possible to assess the main effects of several VASIs (particularly if performance is measured) and to reveal any differences within a manageable number of trials.

An operational trial is less powerful than an experimental trial because the opportunity is greater for extraneous variables to occur randomly and unevenly during the trials, masking the main effects being investigated. Examples of variables beyond the experimenter's control which can influence pilot performance include atmospheric visibility, sun position, wind conditions, other traffic near the airport, radio calls, aircraft functioning, aircraft types, briefing of pilots, etc. These extraneous influences can usually be accounted for, but only by the accumulation of a considerable amount of data (which effectively increase the "power" of the experiment). However, the inordinate amount of effort and expense required to measure flight trajectories by the usual means available (tracking- or kine-theodolites) may limit the number of trials recorded and consequently, the observations may not be statistically representative of either the chosen sample of pilots or the operational parent population. (These difficulties should gradually reduce when improved airborne INS, tracking equipment using radar and laser, and airport facilities with automated data collection like Interscan become more readily available to experimenters.)

Previously, one way of overcoming this measurement problem was to take only one or two points of many flight paths (instead of numerous points in each trajectory) as did Cumming (1962). He used photographs of aircraft crossing the threshold and so could calculate height and ground speed simply for a large number of approaches. Another advantage of his method was the ability to observe covertly the approaches and therefore avoid any unconscious bias in performance (in the main, improvement) which could be present when pilots are aware that they are involved in an evaluation.

The operational trials have the advantage, if a sufficient number of trials is conducted, of providing information about whether the differences found in experimental trials are of practical importance in every-day aviation use. Therefore, operational trials supplement experimental studies, but neither should be substituted for the other.

Further, the results obtained from either of these evaluative techniques cannot subsequently be used to do anything more than predict the suitability of a particular VASI should changes in aviation requirements occur in the future. The effect on human performance of changes in, for instance, aircraft handling characteristics or in approach angles (or other operational requirements) may not become apparent until both an experimental and operational evaluation are repeated under the new conditions. An illustration is provided by the prediction of Cumming (1962) about the quality of the guidance from Red-White VASIS which was mentioned above.

The design details of field evaluations are absolutely critical to the meaningfulness and the predictability of the results. These issues will be considered in more detail further on.

The field trials discussed above have a definite start and finish point (although some researchers may feel that operational trials are never-ending). It is also possible to conduct open-ended field investigations by analysing aviation records as they accumulate.

3.4 Accident and Incident Reports

Probably one of the most obvious ways to the casual observer of checking that a VASI is performing satisfactorily in operation is take note of the number of accidents (or incidents) which can be attributed (either partially or fully) to the VASI. However, such analyses are not straightforward in practice, and relevant information is exceedingly difficult to extract from aviation records. Despite numerous published attempts to analyze accident data (in particular), direct causal relationships remain obscure for many human-factor related events, not only those involved with landing.

It seems that these difficulties arise in part from the subtle complexities of the interactions between man and machine which are generally not well understood. Accidents are seldom caused by one factor; instead frequently quite remote multiple events are causal (Zeller 1970), but these are not easy to identify. Further, even when causes are known, the relevant details may not be recognised or other considerations prevent publication, and these factors compound the difficulty of devising a suitable taxonomy (Wegner 1980). Therefore accidents are often subsumed under the broad headings of "pilot error" or "flight crew error" in the records. Although there have been some attempts to do so in the past, often the categories used are too broad for useful conclusions about causal factors be to reached (e.g. Shuckburgh 1975).

However, in a multi-factorial *post hoc* study of landing accidents, Hartman and Cantrell (1968) were able to pinpoint the presence of an operating landing aid (unspecified) on the runways where most landing accidents occurred. Nevertheless, it remains to be demonstrated whether there is a connection between the landing aid in use and the causal (or instrumental) factors in specific accidents. In one accident there seems to be a clear suspicion that Red-White VASIS was implicated (Millar 1984), but the events occurred so long ago that is it now difficult to establish the interaction with certainty.

Incident data are potentially more advantageous than accident reports because an analysis of trends in this data could indicate possible problems with a VASI and allow corrective action to be taken, thus perhaps avoiding accidents which might otherwise occur. Because accidents are relatively infrequent, substantial amounts of time may elapse before it becomes apparent that a VASI is at fault and requires replacement; incident data could indicate difficulties earlier, hopefully before too many of the suspect VASI type have been installed, increasing the chance that unsatisfactory devices do not remain in service too long.

However, data bases documenting incident reports are more likely to be incomplete than those compiling accident records. This is because the border between what should and should not be brought to the attention of aviation authorities is blurred and some aspects remain within the individual pilot's discretion. It is more likely that the significance of some important events may escape the attention of individuals when clear guidelines for reporting action do not exist, especially if pilots detect overtones of punitive measures connected with the reporting structure. Investigation of causal factors in incidents may also be more difficult as there is no physical evidence (such as a wrecked aircraft fuselage) and important factors may be transient and unlikely to be detected at a later date. Military organisations probably keep the most complete records for analysis, but these usually are not readily available to specialists outside the defence network and the information documented may not be entirely relevant to other aviation categories.

Less formal analyses of incidents obtained from questioning pilots, despite the anecdotal nature, has provided useful information about conditions which make landing aircraft difficult (Lane and Cumming 1959; Armstrong, Bertone and Kahn 1972) and presumably could also indicate whether a VASI is effective operationally. Depending upon the number of pilots and their previous experience, the information from this source could be statistically misleading, and so it could be useful more as indicative of trends worth investigating rather than as a concrete indication of deficiencies (or advantages).

4. FIELD EVALUATIONS: DESIGN AND ANALYSIS OF EXPERIMENTS

The experimental design of any VASI evaluation requires careful execution to avoid the inadvertent introduction of extraneous and misleading influences on the data collected. Suitable experimental design can be quite complicated, depending upon the information and the generality of the conclusions which are ultimately required. In particular, the assignment of subjects to experimental groups, the choice of experimental and comparison conditions (e.g. the aircraft to be flown, the status of the pilots, the number of observations, etc.), and the subsequent data analysis should follow certain well-established principles to avoid undue influences on the data and to ensure that the results are robust (meaning repeatable). In the following section some aspects of the design of field evaluations are discussed. Mostly, examples from experimental field tests have been chosen to illustrate the main principles of experimental design, but many of the same features also apply to operational tests.

4.1 Choice of Aircraft

The choice of aircraft used to flight test VASIs is more critical now than was the case twenty years ago, when most aircraft had piston engines driving propellers. The modern jet engines of large transport aircraft perform differently during descent from the piston engines which predominated in earlier years. Sudden changes in the flight path are usually more difficult to implement from the lower thrust settings used during approach because jet engines have poorer acceleration responses (Davies 1971). In particular, jet-driven aeroplanes without the advantages of extra lift provided by propeller wash have slow response characteristics. Consequently, the flight path of large transport aircraft should be closely monitored so that large deviations in height, speed and rate of descent do not occur (Davies 1971). This is especially true during descent since these aircraft approach at airspeeds below that for minimum power (commonly called "the back side of the drag curve") where dynamic instabilities may be present. Smith (1981) has also pointed out that modern aircraft have higher approach speeds and the inability to extend runway lengths at existing airports has meant that minimum possible approach speeds are selected, making large passenger aircraft more susceptible to the effects of wind shear.

Often, VASIs are tested using aircraft similar to the smaller passenger ferries of the feeder airlines (e.g. F27, King Air, Aero Commander) which are easier to obtain for tests and cheaper to fly, but the power to mass ratio is less than for large airliners. Although visual guidance to assist with maintaining a glideslope is helpful to pilots of these smaller aircraft, faster corrections can be made and descent rates can be arrested more easily. Therefore, conclusions drawn from trials involving smaller aircraft types are generally not suitable for predicting the performance of pilots in larger aircraft.

With the increasing use of STOL aircraft and helicopters in commercial aviation, the potential application of VASIs has increased. These two classes of aircraft require special mention because only sparse experimental work has so far been directed towards guiding their pilots, a factor also recognised by Smith (1981). Fixed-wing STOL aircraft may or may not be stability augmented; each type has obvious differences in handling qualities and perhaps their pilots require different approach guidance commensurate with the handling. Some non-stability-augmented STOL aircraft are difficult to hold accurately on glideslope when the wind conditions are gusty because the controls are not rate damped. Therefore some pilots claim that the change in information from a VASI should not occur as rapidly for these aircraft, but this may be insufficient reason to deprive pilots of information, and the proposition remains to be validated experimentally. According to Whittenburg and his colleagues, helicopters, in contrast to fixed-wing aircraft, have considerable flexibility and do not seem to be so constrained by their flight dynamics and therefore it may be possible to design VASIs giving higher priority to operational requirements (Whittenburg *et al.* 1964).

As many modern aircraft types exhibit quite different response characteristics, it is evident that aircraft representative of the intended user population should be selected for testing VASIs. This does not mean, of course, that ALL aircraft from the population of potential users should be exhaustively tested, but careful selection of aircraft classes could improve the generality of results obtained in experiments.

4.2 Choice of Pilots

The choice of pilots to act as subjects in an experiment is becoming complex since most commercial pilots have been exposed to, at least, the theory of VASI operation or have approached (often frequently) using VASI guidance. In contrast, during the 1950s and the early 1960s, most pilots were unfamiliar with VASI developments and hence experimental evaluations did not need to account for the effect of previous experience on either performance or opinion. It was therefore simpler to assess the relative merits of VASIs under test without needing to consider either pre-existing bias or the influence of previous experience on the interpretation of the signals, flying style or opinions.

Today, the choice of subjects needs to be broadened so that the effects of previous experience on pilots with the new VASI design is revealed over the continuum of likely users; the range must encompass those who are inexperienced and those who have had vast experience. Certain military *helicopter* and general aviation pilots remain a good source (at least in Australia) of pilots who are relatively inexperienced with VASIs, although these pilots are unlikely to have flown many hours in transport aircraft.

The influence of previous experience on flying style has been observed by Lewis and Mertens (1979) who commented that their results from a simulator study indicated that previous experience of Red-White VASIS may have influenced pilot strategy with PAPI. Their pilots were observed to react immediately to the "off-course" signal changes as deviations in position instead of interpreting those signal changes as indicating deviations in rate as Smith and Johnson (1976) intended. Paries (1979) also formed a similar opinion from observing pilot performance during flight trials. Subjective opinions of pilots may also be influenced by their previous experience and not correlate with objective performance parameters. Vaughan, Luce and Kassebaum (1962) discuss an experiment by other investigators in which this was the case; it was not until the pilots had had considerable experience with one lighting pattern that their opinions matched their performance.

Some published reports (e.g. Morrall 1960) have described trials using "in house" pilots who were test pilots at the same establishment where the aid being evaluated was invented and developed. There are two reasons why using the results from experiments involving similarly experienced pilots such as these for predicting how pilots from the commercial community might fly could be misleading. Firstly, special knowledge of one VASI could prejudice flying performance or opinions expressed during an experiment (despite awareness of this possibility by the pilots), and secondly, there is a risk that test pilots represent the upper end of the spectrum of ability amongst the aviator population and their skills may not be possessed by the inexperienced, or indeed the average, pilot. For example, it has been suggested that the strategy of "clicking" the lights of PAPI between red and white (which results in a shallow scalloping profile and is evident in the three trajectories drawn by Smith and Johnson in 1976) demands that the pilot be extremely knowledgeable about the handling characteristics of the aircraft being flown; other less skilled pilots may not be able to do the same (Ross 1980).

There have been reports of differences in flying abilities between various groups of aviators. In one study by Swaroop and Ashworth (1978) landing performance between research and general aviation pilots was compared. They found that research pilots tend to intercept the glideslope at a lower height, fly less steep mean approach angles, and touch down closer to the threshold. Further, the approach techniques of the research pilots appeared to reflect fewer inter-pilot differences than those of general aviation pilots. In another study of instrument approaches (Simmonds 1960), inexperienced pilots flew less consistently from trial to trial (i.e. the intra-pilot variation was large) than those pilots with greater numbers of flying hours. These trends in flying ability could influence evaluation trials considerably. The results from trials using exceptionally skilled pilots might therefore show no differences between VASIs (i.e. the differences were insignificant), but this may not be an accurate reflection of general pilot population performance when using the aids.

4.3 Group Designs

Various procedural details of experiments also have to be carefully planned and here two issues about group designs, whether the same subjects should fly in all experimental conditions and the order of presenting the experimental conditions, have been singled out in connection with transfer of training effects. Transfer effects, which bias the observed magnitude of performance measures between experimental conditions, are common in tracking tasks (Poulton 1974) and these effects probably also occur during flying experiments. For example, simply practicing during the experiment could bias the results. We could expect that performance will improve as the pilots become more familiar with the experimental conditions, particularly if individuals have not approached to the experimental runway before or if the weather is less than optimal. Therefore if VASI A were tested first it would be disadvantaged in a comparison where all the pilots flew to VASI A before VASI B. Traditionally, the way to prevent confounding the results with practice effects is to use an A/B, B/A design, which incorporates half the subjects tracking under Condition A and then under B (the A/B group) while the other half receives the reverse order (B/A). This design will account for symmetric transfer (i.e. each VASI group equally improves throughout the test), which is principally due to practice.

Asymmetric transfer effects on flying style may also be present and can act by differentially influencing each experimental group of subjects, so that the order in which the experiment is conducted becomes critical. For instance, normally pilots may fly characteristic approach paths using different strategies with VASI A and VASI B, but their flying style with either may be altered (or influenced) by a previous requirement in the experiment to fly to the other VASI first. In the least favourable case, performance on VASI A may be extremely variable during the experiment because VASI B induces the pilots to fly erratically with A, but A does not have this effect on B. We may therefore reach the wrong conclusion that VASI A is less suitable, whereas in routine operations it could provide better guidance than VASI B. A separate group design where each subject is assigned to only one condition is usually required to compensate for these effects (Poulton 1974). Such a design was used by Lewis and Mertens (1979), but it is more usual for all subjects to be tested under all of the experimental conditions (e.g. Morrall 1960; Baxter, Cumming, Day and Lane 1960; Paries 1979, etc.) probably for economy.

4.4 Comparison and Control Conditions

In any evaluation a comparison or control condition is required. Unless the proper control is included it is difficult to judge whether the performance of the new VASI is adequate, because it is impossible to conclude with any certainty whether the performance was better than that using another aid or that during unaided approaches. For instance, when assessing the adequacy of theshold clearance it is not sufficient to measure crossing height using only the aid in question without having comparative data (i.e. from a control condition such as "no aid" or another aid). It is possible that a variety of factors could be responsible for pilots achieving clearance during testing and the fact that an aid was installed near the runway may have been merely fortuitous.

The "no aid" condition is sometimes not a suitable control, especially on occasions when the subjects are prejudiced in favour of, or expect that they will perform better with the assistance of a VASI. In this case the Hawthorne Effect (changes in performance stemming from the attention of outside specialists—also see below) may influence performance during an experiment, but may not be sustained indefinitely during normal operations. Therefore, when this sort of reaction is expected, it may be a better procedure to undertake a comparative evaluation instead, where at least two (new) VASIs are tested. It could be possible then to exclude a "no aid" control condition since the contrast would normally be provided by the differences between the two aids. However, caution may be required when interpreting experiments which exclude the "no aid" condition because sometimes the presence of an aid does not substantially improve performance compared with when the aid is absent: this is the case with Red-White VASIS in its 2-bar form (Cumming 1962) and in its simplified version (Smit 1975).

Comparison conditions also need to be chosen so that the VASIs are tested under equivalent conditions, making the comparison fair, but this has not always been the case in previous evaluations. Sometimes VASIs are installed on different sides of the runway (e.g. Brown 1979) or pilots see only one of the VASIs that they are asked to compare during the experiment, relying on their memory for details of the other (e.g. Jones 1977).

4.5 Statistics

As a general rule, the results from evaluations and comparisons need to be subjected to a statistical analysis to determine whether the data shows treatment effects or whether the results might have been expected by chance. The statistical techniques used are properly chosen prior to the beginning of the experiment, but are not applied until the end of the data collection phase. Let us assume for the discussion below that we are analysing the performance parameters from an experiment involving two VASIs with one aircraft type being flown. (These conditions merely simplify the discussion.)

The results are influenced by several factors which contribute to the variance (i.e. the scatter) in pilot performance, and not only those of interest to the experimenter. The measures reflect a combination of intra-pilot variation (e.g. any one pilot may fly several approaches using a different strategy every time), inter-pilot variation (e.g. each pilot may have their own personal flying style), variation on the measures from the experimental conditions (e.g. in the simplest case, VASI A may be different from VASI B) and random (including measurement and extraneous) error. Now the effect that we are interested in is that of the experimental manipulations, i.e. that of global differences between VASIs. However, often the other sources of variation are so large or they combine in such a way that it is difficult to say whether the experimental conditions contributed significantly to the measures if the results are compared only by the ubiquitous "eyeball". (Nonetheless, some effects are large enough to be immediately noticeable in the raw data.) Statistical techniques can be used to sort out the sources of variance and to determine whether there were significant differences between experimental conditions due to the influence of the VASIs, i.e. the experimenter can formulate two hypotheses, H_0: the performances are equal and H_1: the performances are not equal, and determine which to accept.

A statistical technique such as ANOVA (analysis of variance) can be applied to the results to partition the sources of variation (from pilot strategies, both intra- and inter-, the VASIs, and random) and test whether the contribution to the scores from the experimental (or between) conditions is significantly different (i.e. does VASI A outperform VASI B?). Significance in this context refers to a statement of probability about the likelihood of the null hypothesis being true. The experimenter can then decide whether any observed differences were due to chance (or random) events or whether the experimental hypothesis (i.e. H_1) can be reasonably accepted, i.e. if we repeated the experiment would we find the same trends and effects?

Another statistic which is commonly referred to in VASI research is the average or the mean of a set of results. The use of this statistic warrants special mention because it is often quoted (e.g. Paries 1979; Brown 1980; Bisgood, Britton and Ratcliffe 1979) as justification that differences did occur in each condition. However, the mean is not necessarily a reliable indicator on its own. It can be misleading if a measure of the variance in scores around the mean does not accompany it (such as range, root mean square, standard deviation or standard error, etc.). Say that average (or mean) performance with a particular VASI (VASI A) was closer to the nominal glideslope than with another aid (VASI B) undergoing test. It might not be unreasonable to expect, then that A would be better than B in service. However, if the pilots flew more erratically with VASI A than with B (and therefore the variance of the measures with A was larger), approach and landing performance with VASI B may prove to be more reliable in operation.

Further, since the reliability of many statistical tests is dependent upon the distribution underlying the data, if a measure of the shape of the distribution around the mean is not obtained, it could be difficult to select an appropriate test for assessing whether observed differences between mean performance were due to the experimental conditions or to chance events. In most cases a statistical test is required to determine whether any differences exist in the data (unless the distributions are so far apart that they do not overlap much). If the underlying distributions are skewed, the mean may be unsuitable as a description of the "average" performance and the median or mode could well be more appropriate.

Typically, opinions of pilots are presented in reports of evaluations in the form of tabulations or relative frequencies of responses in each category often without further analysis. This data should also be analysed by statistical means to determine whether the frequencies of responses in each category differ significantly. If they do differ then it is possible to say with some certainty whether opinions about the aids under test are distributed unequally amongst the subjects, i.e. do the majority of the participants prefer one aid to another? Without a statistical test, conclusions drawn by merely inspecting the data may be incorrect, especially where small numbers of pilots participate in an experiment. With small samples of opinions there is a strong chance that a random selection of subjects will result in an apparent majority preference for an aid which is not held in the general population.

Statistical analysis can be totally meaningless if the experimental protocol and group design are inappropriately selected to answer the questions posed at the beginning of the evaluation. The succinct adage of "garbage in—garbage out" applies. For instance, interactions between the main effects and, say, asymmetric transfer or prior bias in the subjects cannot usually be accounted for no matter how advanced the statistical analysis. The experimental design and analysis of complicated comparisons is an extremely thorny subject and the reader is referred to statistical texts (Keppel 1971; Winer 1966) for assistance. The above discussion has only just touched the surface of the topic.

4.6 Assessing Pilot Opinion

Pilot opinion may be sought from either operational or experimental evaluations or both. During experimental trials (which usually have small numbers of subjects) it is convenient to use either the structured or unstructured personal interview technique to elicit opinion. The unstructured interview, where the questions are not predetermined, provides an opportunity (if the interviewer is skilled) to obtain answers that might otherwise be missed or inhibited or biased by the form of structured interviews or surveys (Baxter, Cumming, Day and Lane 1960). The structured interview, where the questions are determined prior to the interview allows the experimenter to alter the wording of the questions to suit the understanding of the person being asked, but can produce answers that are easily categorised in much the same way as do written surveys.

During operational trials which usually involve large numbers of subjects, often difficult for the experimenter to contact, interviews are mostly precluded and a written questionnaire must be used instead. A good deal of skill is required when wording written questions so that meaningful answers are obtained. For instance, survey questions should be clear, direct and require only one answer; jargon should be used only if it is readily understood and has a similar meaning to all participants; the wording should be unambiguous without expressing opinions of others. (See Bourchard 1976 for a synopsis and Payne 1951 for a fuller description of many other important features.) Counterbalancing between subjects of the order in which merit categories are listed and/or other procedures designed to overcome response bias and halo effects are also important procedures for determining, for instance, how strongly pilots feel about any particular VASI in a comparison.

It is equally important that the contents of a survey concentrate upon assessing the effectiveness of specific design features contributing to the VASI guidance rather than dwelling on minor or irrelevant aspects (Vaughan, Luce and Kassebaum 1962). Questionnaires which probe known weaknesses of one particular VASI while emphasising strengths of another to the virtual exclusion of determining opinions about general guidance aspects of either, merely provide self-fulfilling prophecies and are not useful indicators of merit.

The results from a structured interview or written survey can be analysed in a relatively straightforward way when prespecified categories of responses are incorporated into the questions. Unstructured interviews usually produce a mass of answers which are difficult to categorise and collate, but there are advantages since a more extensive coverage of the topic could be achieved. Consequently, the preferred experimental technique may be, first to undertake preliminary interviews with a small, but representative, group of the prospective subjects, and then to sample opinions from a larger group. This latter survey may be expedited by a formal structure based on the answers received in the preliminary work.

It is general practice to sample the opinions of pilots who could be required to use the VASI routinely. While no-one would question the validity of doing so, as it is important to be confident that unforeseen barriers do not arise in operational use, this type of evaluation is fraught with problems. Usually, a voluntary postal return system is the only practicable method of obtaining a large sample of opinions. When subjects self-select themselves as is inevitably the case in postal surveys, predictably mostly those with strong motivation to return the survey do so, unless special efforts are made to entice responses. Sometimes only 25 to 30 per cent of the population under question returns answers, and the experimenter is left wondering about the opinions of the remaining 70 per cent who are not represented. In 1962, Cumming found that his postal survey showed an almost equal division of opinion about preference for Red-White VASIS or T-VASIS in an operational environment, yet all previous and subsequent experimental evaluations (Baxter, Cumming, Day and Lane 1960; Hyman 1963; Alexander 1962; Jones 1977; Lewis and Mertens 1979) have shown a marked preference for T-VASIS when opinions are expressed by all the participants.

Cumming (1962) expected a bias in the postal return, *a priori*, and so disregarded his results. In doing so, Cumming implicitly supported the argument that unless a relatively high proportion of the population questioned replied, the experimenter can never be sure on the basis of just one survey whether the sample of opinions obtained reflect the views held by most pilots who used the aids. The experimenter can be more confident about the results if they are compared with opinions obtained in another controlled (statistically representative) sample from the parent population. It is not sufficient to merely repeat the survey, say at another location, because the same factors may be influencing the return rate and similar results may merely confirm the presence of these factors, rather than vindicating the adequacy of the assessment method. The basis used for claiming that PAPI was preferred by the majority of pilots (see Millar 1984) suffers from this shortcoming.

The length of surveys administered in an *ad hoc* fashion must also be minimised because subjects are inclined not to complete all the questions but to concentrate on those of personal interest. This tendency introduces another source of potential bias and statistical analysis, indeed if it is appropriate at all, is very difficult when data is missing.

Despite the care taken in wording and the design of surveys or interviews, it is apparent from some evaluations that pilots can form negative opinions of certain VASIs on the basis of personal bias or previous experience. It has been claimed that national pride can bias subjects in an experiment to prefer the device invented in their own country instead of another device of foreign origin (Whittenburg *et al.* 1964), although previous experience may have been a more important influence in this instance (Vaughan, Luce and Kassebaum 1962). In other circumstances the Hawthorne Effect may be important (this effect can encourage the attitude that "anything new must be better than the old") and so influence pilots to respond positively to an aid that they later think is inferior to the older device. Sometimes subjects even reject a VASI out of hand without further consideration if, perhaps, the equipment does not perform as well as they expect (for instance the light output may be too intense), or if it does not meet one of the requirements that they regard as necessary (for instance portability). Therefore, it is necessary to ensure that opinions are sampled fairly and this may mean in some cases that prototypic equipment should not be used.

5. DISCUSSION

The extent and comprehensiveness of the requirements for an effective VASI may have surprised the uninitiated reader. However, upon reflection, the every-day use by passenger airliners demands that a VASI should satisfy strict criteria. Most of these requirements, particularly the ergonomic principles, can also be directly applied to VASIs for other applications. (Items in the remaining categories, especially the operational requirements, may be altered where appropriate.) As a rule, to ensure that any VASI system does meet the stated criteria, the evaluative procedures and tests also have to be rigorous and extensive.

According to Whittenburg *et al.* (1964) who reviewed various tests of marking and lighting systems extending from 1946 to 1961, omissions or faults in the design of tests, execution and/or data analysis severely limited the information obtained with a consequent waste of valuable resources. There seems to be little in the current literature about VASIs for major airports suggesting that the situation has changed.

There has been a marked tendency for evaluators to rely on the results of usually one, or maybe two, evaluative techniques in isolation when assessing the suitability of a particular VASI.

This reliance can be misleading because, as we have seen, individual evaluative techniques have different aims and therefore provide only part of the mosaic of evidence required. A major concern is the emphasis commonly placed upon operational trials for assessing VASIs to the almost total exclusion of other methods.

Operational trials are the end-point in a VASI development and test schedule and follow on from experimental trials. Laboratory studies (including simulator trials) precede any field trials. Testing of designs in the laboratory can provide information about aspects of the man-VASI interactions that may be missed in the field where it is harder to control the influences of extraneous variables on the measured parameters. In addition, variables may be added or deleted at will and those of importance delineated. There is a risk, however, that laboratory experiments may reveal trivial or irrelevant differences or effects (Chapanis 1967) and these need to be interpreted in the light of field experimentation. Although extrapolation between the depicted and the real world is also a concern with sophisticated manned flight simulators and limits their role, the more realistic characterisation of flying situations can be useful in obtaining designs with potential for development.

Better control of random influences on parameters is obtained in experimental trials than in operational ones and so it is easier to reveal any differences between VASIs. Further, experimental trials may circumvent possible problems arising from prematurely releasing a VASI into an operational environment. Obviously, a VASI should be carefully tested so that it can be assessed whether the device may unexpectedly fail or provide misleading information. Sometimes it seems that the absolute necessity for flight testing new VASIs under an experimental regime has been overlooked in preference to operational evaluations or perhaps the experimental trials are never published. (This survey has not so far revealed any published data on the details and results of developmental or experimental trials of PAPI, PLASI or Red-White VASIS.)

Operational trials serve the necessary purposes of assessing any characteristics which may have been found during experiments under routine conditions and introducing new devices to the members of the intended pilot population. However, often operational evaluations are conducted with the intention of completely assessing the suitability of an aid without considering the findings from other testing techniques. When inadequate sample sizes are collected and statistical analysis is the exception rather than the rule, most operational tests will be unable to predict whether the chance of an accident occurring has been reduced by a VASI. One method of operationally testing a VASI is to use it over a long term, and to then assess its performance by referring to records of landing success. However, this technique is not without its risks, especially if a device is not tested adequately before being introduced into operation. Currently in-service evaluations have rarely been pursued and the methods require refinement, especially in the reporting and recording of incidents and accidents.

Operational trials seem to take precedence over experimental trials, apparently to assess suitability by user-acceptance, but opinions could be conveniently predicted by the judicious choice of pilots who serve as subjects in an experimental field trial. Indeed the majority of field tests, both operational and experimental, have relied principally, if not completely, upon the opinions of pilots to assess the aids being compared. This over-emphasis on the opinions of pilots as the main measure of the successfulness of a VASI design is unwarranted. Pilots cannot always assess their own performance accurately (Vaughan, Luce and Kassebaum 1962) and the influence of previous experience can outweigh the best intentions to forgo or avoid bias. In addition, analysis of the probability of accidents occurring could not be obtained accurately enough to be reliable from subjective data.

The proper use of experimental procedure, design, measurement and statistical analysis of results should enable evaluations to produce reliable conclusions which are not confined to situations containing the vagaries of the experimental conditions. The importance of statistical analysis has been emphasised above, but nevertheless, the use of statistics is perhaps not the panacea that it first seems. Many statistical tests rely on the principle of central tendency which averages data during computation. Therefore the importance of an uncommon event (such as an unacceptably low approach) may be occluded if statistical significance is the only criterion used for assessing results. Further, statistics cannot tell the experimenter whether VASI A is better or worse than VASI B; the level of significance merely indicates the likelihood of finding the same results if the experiment were repeated. The judgement of "better" or "worse" is a value judgement imposed by the experimenter who is responsible for this decision.

An omission in the guidelines outlined above is a discussion of the performance parameters that could be collected during an assessment. Here again the choice of specific parameters is the responsibility of the experimenter who must take into account the aims which have motivated the trials. For instance, some trials ascertain whether adequate threshold clearance will be achieved for certification purposes, and so it is expedient only to record height above threshold to predict the sufficiency of the guidance so long as figures for comparison from unaided or other approaches are available. On other occasions fuller information about the entire approach may be required.

Root mean squared (RMS) error (in height or angle) around the nominal glideslope is a metric which is often used to evaluate performance, but it is on no account the ONLY useful measure, nor even the most appropriate for all tests. During a recent field experiment by Anderson, Selway and Millar, an airborne observer (experienced helicopter pilot) judged that the experimental helicopter pilots were making large control inputs in response to the flashing signals from PLASI. It is quite probable that the pilots were actually flying reasonably straight trajectories, but overcontrolling a helicopter is potentially dangerous since, amongst other situations, tail rotor authority may be lost. Because the control movements of the pilots were not measured during these flights the accuracy of the comments could not be directly confirmed and it is unlikely that such features would be apparent in the RMS values.

The available measuring equipment has influenced the actual choice of parameters for measurement in previous experiments also. As we have seen, Cumming was able to photograph sparingly a few points in each trajectory of many aircraft on approach and yet still successfully differentiate between VASIs in a meaningful manner. Others, (e.g. van Oosterom 1963), have recommended that a multitude of parameters should be taken, including the pilot's heart rate, the number of control movements (for distinguishing effort expended) as well as the more usual measures of performance to estimate accuracy in maintaining the glidepath.

Van Oosterom did so because he felt that it should be possible to assess a VASI in only one set of trials, thereby obviating the need for several tests by scoring many parameters. (Incidentally, van Oosterom's scoring method is recognised in zoology under the title of "numerical taxonomy" and has been used for classifying evolutionary relationships between species. The difficulty in the method, which presumably holds for VASI evaluations also, relates to the weight that any one parameter should receive in relation to another.) A similar theme has also been echoed by H. J. Clark (1968) who felt that the lack of a standard assessment technique inhibited the usefulness of VASI tests. Given the variety of possible evaluative techniques, however, it seems that only broad procedural outlines relating to solid experimental design could possibly be adopted for the various tests.

Currently, the objective performance measures used for distinguishing between VASIs describe the movements of the aircraft rather than the efforts of the pilots to achieve this performance. It was considered by Vaughan, Luce and Kassebaum (1962) that measuring aircraft performance was probably too indirect for assessing pilot performance rigorously enough because in many experiments differences between various marking and lighting patterns were not exposed by these measures. They hypothesised that either any differences were too minor to be reflected in the measures or that pilots were able to compensate for differences by adopting appropriate flying strategies. If differences do in fact exist then it may be possible to differentiate between aids by characterising the effort pilots make to acquire their flight paths by using workload measures as the dependent variables in conjunction with the more traditional parameters describing flight paths, as these earlier authors have suggested.

There seems to have been minimal amount of research effort devoted lately towards understanding the fundamental aspects of good VASI design for most applications (with the exception of research about carrier landings in the USA—see Kaul, Collyer and Lintern 1980 for a partial review); rather, most trials concentrate on answering particular queries, often obtaining little more than information about the efficacy of specific hardware. This may be one of the reasons why we find VASIs currently being developed and tested which incorporate the same or similar guidance concepts as those of previous designs, many of which have undesirable characteristics.

Examples of the repetitive inclusion of similar guidance concepts in different VASIs abound. For instance, both PAPI and PLASI incorporate colour-coding despite many recommendations based on operational observations and careful laboratory experiments that it is dangerous (see Clark and Gordon 1981 for a review of these arguments). Both PAPI and PLASI also have point-source origins of their signals which expresses error from the nominal glidepath in angular terms. Sparke (1958) criticised angular feedback in connection with the PVG, claiming that pilots require feedback in terms of height from the nominal glideslope. Sparke's opinion was based on theoretical grounds rather than from an established experimental basis and the correctness or otherwise of his view remains to be validated, despite the continued use of angular signals in a multitude of different VASIs. In addition, the objective lenses which emit the signal in PAPI boxes are not adequately protected from the environment and build-up of dirt or moisture on the lenses can degrade or destroy the signal as happens in a variety of other VASIs which are designed along similar lines. Further, PLASI also uses flash coding and the pilot is required to notice the temporal relationships between the on and off periods to determine his deviation from glidepath; a coding concept that has been rejected in railway signalling because the human visual system is unable to judge temporal patterns consistently (Mashour 1974).

Basic experimental work on the fundamental aspects of guidance that a pilot might require can reap substantial improvements. Recently, Kaul, Collyer and Lintern (1980) demonstrated in a simulation of carrier landings that adding descent rate, especially if command signals are also provided, can significantly improve the accuracy with which pilots track the glideslope using a Fresnel Lens Optical Landing System. As yet these improvements have not been completely flight-tested and so can not be considered as a general rule for guidance principles. However, their "command" condition closely resembled a quickened display which Jensen (1981) and others have shown does markedly improve tracking in experiments. The influence of command information on performance was suspected many years ago by Cumming (1962) and his associates based upon well-known ergonomic principles again derived from laboratory experiments, but commands have been included in only one ICAO-approved VASI, T-VASIS, so far.

Zeller (1972) predicted that landing accidents would become less common relative to other classes of accidents if attention were devoted to improving landing and navigational aids. The observations of Hartman and Cantrell (1968) where they stated that a landing aid of some sort was in operation at the time most landing accidents occurred (in the general aviation sphere of the USA, at least) suggests that perhaps further effort could be profitably expended. Although Hartman and Cantrell did not state whether these landing aids were VASIs or other types (e.g. ILS), Red-White VASIS is extensively installed in the USA and was probably operating on a significant number of occasions when accidents occurred, and so their statement warrants further exploration. Zeller (1981) has expressed the opinion that to prevent accident rates rising, known principles of training and operation must be assiduously applied, and so it could be expected that further improvements in landing success may be achieved by replacing the less effective of the VASIs currently being used with better designs.

6. CONCLUSIONS

Visual Approach Slope Indicators are required, in essence, to REPLACE the external visual world to assist the pilot to perceive the glidepath accurately, consequently reducing the risk of a landing mishap. Therefore a VASI certified for international operations should substantially fulfil the numerous conditions which have been listed above for use by large transport aircraft. Accordingly, evaluative tests are also required to be exhaustive and to adhere to the rules and procedures that govern experimental design. Several examples of where poor technique or inadequate experimental design could, or did, bias conclusions have been drawn upon to illustrate this point.

The application of reasoned experimental design to the evaluative tests conducted on VASIs could provide more reliable and (more) complete information than is currently the case. The unnecessary concentration upon operational trials evident in many evaluations redirects available resources of money and effort away from the potentially more informative experimental trials.

Most operational trials introduce a specific VASI to the general user population and collect only opinions about acceptability, risking at the same time biasing the data by the unpredictable effects of prior pilot experience. Frequently insufficient numbers of opinions are sampled and statistical tests to ascertain whether these are representative of the entire population are either not performed or are inappropriate. The information is usually obtained from postal surveys and is likely to be too abbreviated to predict accurately whether the aid will operate as intended in service. Further, many of these operational trials do not include a control or comparison condition and therefore restrict the meaningfulness of the results. All of these shortcomings make it difficult to choose between competitors for the same users. Hence various attempts at standardization suffer.

Relying on pilot opinion can be misleading when field trials of any sort are being conducted, and so it is recommended that objective measures take precedence. One advantage of objective data lies in the possibility of calculating whether a VASI reduces the probability of an accident compared to the incidence of normal unaided approaches or to that obtained with another aid. However useful this statistic may seem, it is seldom calculated. It could be profitable to investigate parameters that might be useful indicators of pilot strategies and workload. These measures when used in conjunction with conventional measures of aircraft flight paths might distinguish more finely between VASIs during experiments and therefore provide a better idea of VASI performance under operational conditions.

Future development of VASIs is hindered by the narrow objectives of operational trials and substantial improvements in new VASI designs could be expected if more attention were given to other means of assessment, especially experimental techniques. The same principles and conclusions apply to VASIs developed for other specific purposes including those of the military.

ACKNOWLEDGEMENTS

Drs K. W. Anderson and B. A. J. Clark provided many useful comments on previous drafts of this report and Mr A. Ross vetted the final version. Mrs J. Gordon assisted with collating the literature.

REFERENCES

AGARD (1967). Aircraft instrument and cockpit lighting by red or white light. Conference Proceedings No. 26, Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organisation, Symposium 30–31 October, Belgium.

AGARD (1980). Fidelity of simulation for pilot training. *AGARD-AR-159*, Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organisation, France.

AGARD (1981). Characteristics of flight simulator visual systems. *AGARD-AR-146*, Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organisation, France.

Alexander, R. A. (1962). A RAAF evaluation of two visual glidepath systems. *HE Note 12*, Aeronautical Research Laboratories, Melbourne, Australia.

Armstrong, G., Bertone, C., and Kahn, O. (1972). Study of optical illusions during visual approaches. *FAA-RD-72-146*, Department of Transportation, Federal Aviation Administration, Washington, D.C., U.S.A.

Baxter, J. R., Cumming, R. W., Day, R. H., and Lane, J. C. (1960). A comparison of three visual glidepath systems. *HE Note 8*, Aeronautical Research Laboratories, Melbourne, Australia.

Baxter, J. R., Day, R. H., and Lane, J. C. (1960). The sensitivity of the Precision Visual Glidepath (P.V.G.) at long range. *HE Note 5*, Aeronautical Research Laboratories, Melbourne, Australia.

Baxter, J. R., and Lane, J. C. (1960). The "Tee" Visual Glidepath (T.V.G.). An alternative type of visual approach aid. *HE Note 7*, Aeronautical Research Laboratories, Melbourne, Australia.

Billings, C. E., Gerke, R. J., and Wick, R. L. (1975). Comparisons of pilot performance in simulated and actual flight. *Aviation, Space, and Environmental Medicine*, *46*(*3*), 304–308.

Bisgood, P. L., Britton, J. W., and Ratcliffe, J. Y. (1979). Windshear encounters during visual approaches at night. A piloted simulator study. *RAE Technical Report 79126*, Royal Aircraft Establishment, Farnborough, Hants, U.K.

Bouchard, T. J. (1976). Field research methods: interviewing, questionnaires, participant observation, systematic observation, unobtrusive measures. *In* Dunnette, M. D. (ed.) (1976). *Handbook of Industrial and Organizational Psychology*. Rand McNally College Publishing Company, Chicago, USA.

Brown, M. A. (1978). Agenda Item 9: Future activities—Precision Approach Path Indicator. *VAP/8-WP/10*, Visual Aids Panel, ICAO Eighth Meeting, Montreal.

Brown, M. A. (1979). New Precision Approach Path Indicator undergoing operational evaluation. *ICAO Bulletin*, *34*(*1*), 29–31.

Brown, M. A. (1980). Preliminary draft report of Visual Aids Panel Working Group on PAPI. Civil Aviation Authority, London, U.K.

Buckland, G. H., Munroe, E. G., and Mehrer, K. I. (1980). Flight simulator runway visual textural cues for landing. *AFHRL-TR-79-81*, Brooks Air Force Base, Texas, U.S.A.

CAA (1978). The Precision Approach Path Indicator (PAPI). *CATC 10-CP/6*, Paper by United Kingdom Civil Aviation Authority, Commonwealth Air Transport Council, Tenth Meeting, London, U.K.

Chapanis, A. (1967). The relevance of laboratory studies to practical situations. *Ergonomics*, *10*(*5*), 557–577.

Chase, W. D. (1971). Evaluation of several TV display systems for visual simulation of the landing approach. *NASA TN D-6274*, National Aeronautics and Space Administration, Washington, D.C., U.S.A.

Clark, B. A. J., and Gordon, J. E. (1981). Hazards of colour coding in Visual Approach Slope Indicators. *Systems Report 25*, Aeronautical Research Laboratories, Melbourne, Australia.

Clark, H. J. (1968). A survey of aircraft Visual Approach Slope Indicators. *AMRL-TR-68-30*, Aerospace Medical Research Laboratories, Ohio, U.S.A.

Crymble, C. (1977). Visual illusions on landing. *DCIEM-TR-77X29*, Defence and Civil Institute of Environmental Medicine, Canada.

Cumming, R. W. (1962). Interim report on the operational evaluation of two visual glidepath systems. *HE Technical Memorandum 5*, Aeronautical Research Laboratories, Melbourne, Australia.

Cumming, R. W., and Lane, J. C. (1957). Progress report of Australian research on visual problems of approach to landing. *A.A.R.C. 59*, Australian Aeronautical Research Committee, Department of Supply, Australia.

Davies, D. P. (1971). *Handling the Big Jets*. Air Registration Board, Redhill, U.K.

Dawson, R. D. (1979). The horizontal approach path indicator (HAPI) stable platform. *ASWE RE 79009*, Admiralty Surface Weapons Establishment, Portsmouth, Hants, U.K. Restricted.

DeVore, G. (1981). Personal communication to the author. July, Melbourne, Australia.

DeVore, G. (1982). PLASI—a new concept in visual approach systems. *ICAO Bulletin*, *37*(*4*), 13–16.

Fischer, E., Haines, R. F., and Price, T. A. (1980). Cognitive issues in head-up displays. *NASA TP-1711*, National Aeronautics and Space Administration, U.S.A.

Foushee, H. C. (1981). The role of communications, socio-psychological, and personality factors in the maintenance of crew coordination. Paper presented to the *First Symposium on Aviation Psychology*, *APL-1-81*, Aviation Psychology Laboratory, Ohio State University, Ohio, U.S.A.

Gates, R. F. (1970). Visual Approach Slope Indicator (VASI) system for long-bodied aircraft. *FAA-NA-70-59*, Federal Aviation Agency, Washington D.C., U.S.A.

Gibson, J. J. (1950). *The Perception of the Visual World.* Houghton-Mifflin, New York, U.S.A.

Gibson, J. J., Olum, P., and Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology*, *68*, 372–385.

Gregson, R. E. (1978). Evaluation of Reduced T System. *Report VA-110*, Department of Transport, Melbourne, Australia.

Hald, O. (1980). *Evaluation of PAPI by Denmark*. Personal letter to M. A. Brown, Rapporteur, VAP Working Group on PAPI, from H. Dahl, Directorate of Civil Aviation, Denmark, 12 June 1980.

Hartman, B. O., and Cantrell, G. K. (1968). Psychological factors in landing short accidents. *Flight Safety*, *2*, 26–32.

Hockberg, J., and Smith, O. (1955). Landing strip markings and the 'explosion pattern'. 1. Program preliminary analyses and apparatus. *Perceptual and Motor Skills*, *5*, 81–92.

Hunt, D. P. (1961). The effect of the precision of information of feedback on human tracking performance. *Human Factors*, *3*, 77–85.

Hyman, M. L. (1963). Comparative evaluation of Australian T.V.G. and United States standard visual approach slope indicators. *FAA Project No. 421-2V*, Federal Aviation Agency, Washington, D.C., U.S.A.

Iwataki, N. (1973). Visual problems concerning landing accidents. *NASA TT-F-15.054*, National Aeronautics and Space Administration, Washington, D.C., U.S.A.

ICAO Annex 14 (undated). International Civil Aviation Organisation, Montreal.

Jensen, R. S. (1981). Prediction and quickening in perspective flight displays for curved landing approaches. *Human Factors, 23(3)*, 355–363.

Johnson, N. A., and Poston, A. M. (1976). A comparison of red and white cockpit lighting under quasi-operational conditions. *HEL-TM-6-76*, Human Engineering Laboratories, Aberdeen Proving Ground, U.S.A.

Johnston, I. R., White, G. R., and Cumming, R. W. (1973). The role of optical expansion patterns in locomotor control. *American Journal of Psychology, 86(2)*, 311–324.

Jones, P. M. (1977). VASI improvement—T-VASI Evaluation. *NAFEC Technical Letter Report NA-77-45-LR*, National Aviation Facilities Experimental Centre, Atlantic City, U.S.A.

Kaul, C. E., Collyer, S. C., and Lintern, G. Glideslope descent-rate cuing to aid carrier landings. *NAVTRAEQUIPCEN IH-322*, Naval Training Equipment Center, Orlando, Florida, U.S.A.

Keppel, G. (1971). *Design and Analysis: a Researcher's Handbook*. Prentice-Hall Inc., New Jersey, U.S.A.

Kraft, C. L., and Elworth, C. L. (1968). How high is up? *Interceptor, 10(10)*, 4–14.

Kraft, C. L. (1969). Measurement of height and distance information—provided pilots by extra-cockpit visual scene. In *Visual factors in transportation systems*. Proceedings of Spring meeting, Committee on Vision, National Academy of Sciences, National Research Council, Washington, D.C., U.S.A.

Kraft, C. L. (1977). Windshield quality and pilot performance. *AMRL-TR-77-39*. Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio, USA.

Lane, J. C., and Cumming, R. W. (1956). The role of visual cues in final approach to landing. *HE Note 1*, Aeronautical Research Laboratories, Melbourne, Australia.

Lane, J. C., and Cumming, R. W. (1959). Pilot opinions and practices on the approach to landing: a questionary survey among Australian civil and military pilots. *HE Report 1*, Aeronautical Research Laboratories, Melbourne, Australia.

Lewis, M. F., and Mertens, H. W. (1979). Pilot performance during simulated approaches and landings made with various computer generated visual glidepath indicators. *FAA-AM-79-4*, Federal Aviation Agency, Washington, D.C., U.S.A.

Mashour, M. (1974). *Human Factors in Signalling Systems: Specific Applications to Railway Signalling*. John Wiley and Sons, New York, U.S.A.

Mertens, H. W. (1978). Comparison of the visual perception of a runway model in pilots and nonpilots during simulated night landing approaches. *Aviation, Space, and Environmental Medicine, 49(9)*, 1044–1055.

Mertens, H. W. (1981). Perception of runway image shape and approach angle magnitude by pilots in simulated night landing approaches. *Aviation, Space, and Environmental Medicine, 52(7)*, 373–386.

Mertens, H. W., and Lewis, H. F. (1982). Effect of different runway sizes on pilot performance during simulated night landing approaches. *Aviation, Space, and Environmental Medicine, 53(5)*, 463–471.

Millar, Jane. (1984). An analytical comparison of three Visual Approach Slope Indicators: VASIS, T-VASIS and PAPI. *Systems Report* 33, in publication, Aeronautical Research Laboratories, Melbourne, Australia.

Millar, J., and Selway, R. E. (1981). Information content of Visual Approach Slope Indicators (VASIs). *Proceedings of the 18th Annual Conference of the Ergonomics Society of Australia and New Zealand*, November, Canberra, Australia.

Morrall, J. C. (1960). A flight assessment of the methods of indicating a visual glide path—the R.A.E. Visual Glide Path Indicator and the A.R.L. Double Bar Ground Aid. *RAE Technical Note No. BL48*, Royal Aircraft Establishment, Bedford, U.K.

Murch, G. M. (1973). *Visual and Auditory Perception*. The Bobbs-Merrill Company, Inc., New York, U.S.A.

Owen, D. H., and Jensen, R. S. (1981). Methodological approaches to identifying relevant features for visual flight simulation. *AFOSR TR-81-0479*, Air Force Office of Scientific Research, Bolling Air Force Base, U.S.A.

Paries, J. (1979). Expérimentation de l'indicateur visuel de pente d'approche PAPI (Precision Approach Path Indicator). *Rapport d'Etude No. 223*, Service Technique de la Navigation Aérienne, Paris, France.

Payne, S. L. (1955). *The Art of Asking Questions*. Princeton University Press, New Jersey, U.S.A.

Pitts, D. G. (1967). Visual illusions and aircraft accidents. *SAM-TR-67-28*, United States Air Force School of Aerospace Medicine, Brooks Air Force Base, Texas, U.S.A.

Poulton, E. C. (1974). *Tracking Skill and Manual Control*. Academic Press, New York, U.S.A.

Regan, D. (1981). Visual guidance of a man/vehicle system in a ground or close to ground environment: a literature review and evaluation. Eastern Scientific Consultants, Halifax, N.S., Canada.

Riordan, R. H. (1974). Monocular visual cues and space perception during the approach to landing. *Aerospace Medicine*, *45*, 766–771.

Roman, J. (1977). Establishment criteria for Visual Approach Slope Indicators (VASI). *FAA-ASP-76-2*, Federal Aviation Administration, Washington, U.S.A.

Roscoe, S. N. (1979). When day is done and shadows fall, we miss the airport most of all. *Human Factors*, *21*, 721–731.

Roscoe, S. N. (1980). *Aviation Psychology*. The Iowa State University Press, Ames, Iowa, U.S.A.

Roscoe, S. N. (1981). Landing aircraft, detecting traffic and the dark focus. Paper presented by R. Hennessy to the *First Symposium on Aviation Psychology*, *APL-1-81*, Aviation Psychology, Ohio State University, Ohio, U.S.A.

Ross, A. (1980). Personal communication to the author.

Sams, R. (1980). Ducking under—performance and prevention. *Flight Safety Focus*, *2*, 6–11.

Shuckburgh, J. S. (1975). Accident statistics and the human-factor element. *Aviation, Space, and Environmental Medicine*, *46*(*1*), 76–79.

Simmonds, D. C. V. (1960). An investigation of pilot skill in an instrument flying task. *Ergonomics*, *3*, 249–253.

Smit, J. (1975). Evaluation of an abbreviated visual approach slope indicator (A-VASIS) for use on short, unpaved runways. *NLR TR75066U*, Netherlands Civil Aviation Department, Holland.

Smith, A. J. (1981). Airfield visual aids research at the Royal Aircraft Establishment. *RAE-F3-431, Technical Memorandum*, Royal Aircraft Establishment, Bedford, U.K.

Smith, A. J., and Johnson, D. (1973). An assessment of the in-service performance of VASI. *Technical Memorandum Avionics 131*, Royal Aircraft Establishment, Farnborough, U.K.

Smith, A. J., and Johnson, D. (1976). The Precision Approach Path Indicator—PAPI. *RAE Technical Report 76123*, Royal Aircraft Establishment, Farnborough, U.K.

Sparke, J. W. (1958). Methods of indicating a glide path by visual means. *RAE Technical Note EL160*, Royal Aircraft Establishment, Farnborough, U.K.

Stout, C. L., and Stephens, W. A. (1975). Results of simulator experimentation for approach and landing safety. *Douglas Paper 6395A*, Douglas Aircraft Company, U.S.A.

Swaroop, R., and Ashworth, G. R. (1978). An analysis of flight data from aircraft landings with and without the aid of a painted diamond on the same runway. *NASA CR-143849*, National Aeronautics and Space Administration, Dryden Research Center, California, U.S.A.

Van Oosterom, T. (1963). A flight test method for the evaluation of approach and runway lighting effectiveness. *NATO Report 431*, Advisory Group for Aeronautical Research and Development, North Atlantic Treaty Organization, Paris.

Vaughan, W. S., Jr, Luce, T. S., and Kassebaum, R. G. (1962). Airport marking and lighting systems: a survey of operational tests and human factors, 1959–1961. *HSR-RR-61/13-MKX*, Human Sciences Research, Inc., Arlington, U.S.A.

Vaughan, W. S., Jr, Rollins, W. F., and Luce, T. S. (1963). Laboratory studies of the ability of observers to perform three visual tasks required of pilots during landing. *HSR-RR-63/7-Mk-X*, Human Sciences Research, Inc., Arlington, U.S.A.

Vinache, W. E. (1947). Illusions experienced by aircraft pilots while flying. *Journal of Aviation Medicine*, *18*, 308–325.

Watters, D. L., Rollins, W. F., Frey, R. B., and Cavonius, C. R. (1964). Flight analysis of approach and landing guidance elements of heliport lighting patterns. *HSR-RR-64/7-Mk-X*, Human Sciences Research, Inc., McLean, Virginia, U.S.A.

Wegner, K. W. (1980). A tentative taxonomy of human interactive factors in aircraft mishaps. *AFDSR-TR-80-0595*, Boston College, Massachusetts, U.S.A.

Wewerinke, P. H. (1978). The visual scene perception process involved in the manual approach. *NLR-TR-78130-U*, National Aerospace Laboratory, Amsterdam, Holland.

Whittenburg, J. A., Vaughan, W. S., Jr, Havron, M. D., and Cavonius, C. R. (1964). Airport/Heliport marking and lighting systems: a summary report on human factors research. *HSR-RR-64/2-Mk-X*, Human Sciences Research, Inc., McLean, Virginia, U.S.A.

Winer, B. J. (1971). *Statistical Principles in Experimental Design*. 2nd edition, McGraw-Hill, Kogatusha, Tokyo, Japan.

Winterberg, R. P., Brictson, C. A., and Wulfeck, J. W. (1966). A rationale for evaluating visual landing aids: night carrier recovery. Dunlap and Associates, Inc., Santa Monica, U.S.A.

Zeller, A. F. (1970). Accidents and Safety. *In* de Greene, K. B. (ed.), *Systems Psychology*. McGraw-Hill, Inc., U.S.A.

Zeller, A. F. (1972). Human error in the seventies. *Aerospace Medicine*, *43*(*2*), 492–505.

Zeller, A. F. (1981). Human error in the seventies—reviewed and projected through the eighties· *Aviation, Space, and Environmental Medicine*, *52*(*4*), 241–246.

DISTRIBUTION

AUSTRALIA

DEPARTMENT OF DEFENCE

Central Office

Chief Defence Scientist
Deputy Chief Defence Scientist
Superintendent, Science and Program Administration
Controller, External Relations, Projects and Analytical Studies } (1 copy)
Defence Science Adviser (U.K.) (Doc. Data sheet only)
Counsellor, Defence Science (U.S.A.) (Doc. Data sheet only)
Defence Science Representative (Bangkok)
Defence Central Library
Document Exchange Centre, D.I.S.B. (18 copies)
Joint Intelligence Organisation
Librarian H Block, Victoria Barracks, Melbourne
Director General—Army Development (NSO) (4 copies)
Defence Industry and Material Policy, FAS

Aeronautical Research Laboratories

Director
Library
Superintendent—Systems
Divisional File—Systems
Author: Jane Millar

Materials Research Laboratories

Director/Library

Defence Research Centre

Library

RAN Research Laboratory

Library

Navy Office

Navy Scientific Adviser
RAN Aircraft Maintenance and Flight Trials Unit
Directorate of Naval Aircraft Engineering
Directorate of Naval Aviation Policy
Directorate of Naval Ship Design
Directorate of Naval Safety
Directorate of Psychology—Navy

Army Office

Scientific Adviser—Army
Director of Aviation—Army
OC Australian Army Aviation Centre, Oakey

Air Force Office

Air Force Scientific Adviser
Aircraft Research and Development Unit

Scientific Flight Group
Library
Director General Aircraft Engineering—Air Force
Director General Operational Requirements—Air Force
Director, Operational Requirements B—Air Force
Director of Air Force Safety
Director of Air Force Medicine
Director of Psychology—Air Force
HQ Operational Command
HQ Support Command (SO Psychology)
RAAF Academy, Point Cook
CO, Institute of Aviation Medicine, Point Cook

Central Studies Establishment
Information Centre

Government Aircraft Factories
Manager
Library

DEPARTMENT OF AVIATION

Library
Assistant Director Operations Victoria/Tasmania Regional Office
Flying Operations Section
First Assistant Secretary, Flying Operations
Aviation Medicine Branch
Regular Public Transport
First Assistant Secretary, Airways Division
Visual Aids Section
Advanced Techniques Section
Bureau of Air Safety Investigations, Dr R. Lee
Inspector Air Safety Investigations
New South Wales Regional Office
Queensland Regional Office
South Australia/Northern Territory Regional Office
Victoria/Tasmania Regional Office
Western Australia Regional Office
Assistant Director Airways Branch

STATUTORY AND STATE AUTHORITIES AND INDUSTRY

AWA, Mr R. K. Smith
Kruger Schwartz Pty Ltd, Mr B. Langton

CANADA

CAARC Coordinator, Human Engineering
International Civil Aviation Organization, Library
Defence and Civil Institute of Environmental Medicine
NRC
Aeronautical & Mechanical Engineering Library

Universities and Colleges

Alberta Dr T. M. Caelli

FRANCE

ONERA, Library
Ministere des Transports, Mr J. Paries

INDIA

CAARC Coordinator, Human Engineering
National Aeronautical Laboratory, Information Centre

JAPAN

Institute of Space and Astronautical Science, Library

NETHERLANDS

National Aerospace Laboratory (NLR), Library

NEW ZEALAND

Defence Scientific Establishment, Library
Transport Ministry, Airworthiness Branch, Library
OC, Defence Environmental Medical Unit
CAARC Coordinator, Human Engineering
DSIR., Library

UNITED KINGDOM

CAARC Coordinator, Human Engineering
Royal Aircraft Establishment, Farnborough, Library
Civil Aviation Authority, Directorate of Aerodrome Standards
Commonwealth Air Transport Council Secretariat
British Library, Lending Division
Aircraft Research Association, Library

UNITED STATES OF AMERICA

NASA
Scientific and Technical Information Facility
Lyndon B. Johnson Space Center, Dr G. W. S. Abbey
Department of the Air Force, Dr H. J. Clark
Federal Aviation Administration
Ames Research Center, California
Office of Aviation Medicine, Washington
Aircraft Development Service, Washington
Civil Aeromedical Institute, Oklahoma
National Aviation Facilities Experimental Center, Library
Mike Monoroney Aeronautical Center, Dr H. W. Mertens
Douglas Aircraft Company (Operational Flight Safety), J. S. Clausel

Universities and Colleges

Illinois at Urbana-Champaign	Department of Psychology Institute of Aviation
Arizona State	College of Engineering and Applied Sciences
Ohio State University	Dr R. S. Jensen
New Mexico	Prof. S. Roscoe
Massachusetts Institute of Technology	M.I.T. Libraries

SPARES (30 copies)

TOTAL (145 copies)

Department of Defence

DOCUMENT CONTROL DATA

I. a. AR No.	I. b. Establishment No.	2. Document Date	3. Task No.
AR-003-962	ARL-SYS-R-32	September 1984	DST 82/080

4. Title	5. Security	6. No. Pages
GUIDELINES FOR THE EVALUATION OF VISUAL APPROACH SLOPE INDICATORS	a. document Unclassified	31
	b. title U c. abstract U	7. No. Refs 95

8. Author(s)	9. Downgrading Instructions
Jane Millar	—

10. Corporate Author and Address	11. Authority (*as appropriate*)
Aeronautical Research Laboratories P.O. Box 4331, Melbourne, Victoria, 3001	a. Sponsor b. Security c. Downgrading d. Approval —

12. Secondary Distribution (*of this document*)

Approved for public release

Overseas enquirers outside stated limitations should be referred through ASDIS, Defence Information Services Branch, Department of Defence, Campbell Park, CANBERRA, ACT, 2601.

13. a. This document may be ANNOUNCED in catalogues and awareness services available to . . .

No limitations

13. b. Citation for other purposes (*i.e. casual announcement*) may be (*select*) unrestricted (or) as for 13 a.

14. Descriptors		15. COSATI Group
Visual aids	Visual Approach Slope Indicators	05050
Approach indicators		01050
Landing aids		
Aircraft landing		
Experimental design		

16. Abstract

This report discusses techniques for evaluating the efficacy of Visual Approach Slope Indicators (VASIs) intended for transport aircraft landing at major runways. The discussion is predicated on the assumption that VASI signals should act as a substitute for the information the pilot sees in the natural scene, and a list of design requirements is developed accordingly. Several testing techniques and the utility of information likely to be obtained from each one is outlined. Various aspects of experimental design including the choice of aircraft, pilots, treatments and statistical analysis which can influence the validity of conclusions are highlighted. In many published evaluations to date there has been undue reliance on the subjective opinions of pilots instead of measurements describing flight paths in objective terms. Similarly, operational evaluations have been preferred to more informative experimental tests. Adequate testing of current VASIs could enable substantial improvements to be incorporated into VASI designs of the future, but unfortunately, previous tests have not always been adequate, restricting the

This page is to be used to record information which is required by the Establishment for its own use but which will not be added to the DISTIS data base unless specifically requested.

16. Abstract (Contd)

pertinence of the findings and limiting the generality of conclusions which might otherwise have influenced subsequent designs.

17. Imprint

Aeronautical Research Laboratories, Melbourne

18. Document Series and Number	19. Cost Code	20. Type of Report and Period Covered
Systems Report 32	732265	—

21. Computer Programs Used

—

22. Establishment File Ref(s)

M2/793

The frequency of this series is irregular and not all parts will be supplied. Commencement date for this series was ....1951........ . For your records — if applicable — Defence Research Centre Salisbury was known as Weapons Research Establishment.

END

FILMED

11-85

DTIC