## A Reference Model of Soldier Attention and Behavior

MAJ Jonathan K. Alt
TRADOC Analysis Center – Monterey
Naval Postgraduate School
P.O. Box 8695
Monterey, CA 93943
831-656-3732
jonathan.alt@us.army.mil

Christian J. Darken

Dept. of Computer Science and MOVES Institute

Naval Postgraduate School

Monterey, CA 93943

831-656-2095

cidarken@nps.edu

ABSTRACT: This research expands entity level representation of situation awareness and behavior. Building on previous work, the researchers developed an integrated and tractable modeling framework for the representation of a soldier's visual attention, situation awareness, and target acquisition prior to a close range encounter, as well as the soldier's initial action choice. The researchers developed data to populate the model through subject matter expert questionnaires and a live virtual experiment. The resulting algorithms provide insight into soldier action choice on contact. These and other aspects of the model have been coded for demonstration as a proof of principle, while work on the full reference model continues.

### 1. Introduction

The main contribution of this work is to present an integrated conceptual model of a soldier's visual attention (search), situation awareness, and target acquisition leading up to a close range encounter, and his initial course of action once the encounter begins. The situation awareness representation as a probability distribution of threat location and the driving of search by situation awareness are new to military simulations. While the model as a whole is conceptual, large pieces of it have been implemented in software. The representation of entity level situation awareness in combat simulations is a key capability to informing analysis of future combat systems. These future systems rely heavily on networked capabilities and the concept of a common operating picture. Without intelligent entities capable of perceiving and acting on the information passed through the network, the analysis of the impact of future capabilities on combat operations will be suspect.

The term situational awareness first came into use among the aviation community as early as World War I. The term is now commonly used

throughout most branches of service. The meaning of the term, however, is still ill-defined and subject to much debate. Endsley defines situational awareness (SA) as "the perception of elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future." This definition of SA leads to a three level model of SA (Endsley, M. R., 2001).

The first level of this model is perception of the environment. For the dismounted infantry soldier this translates to everything he can sense with his five organic senses. The second level of this model entails the individual's reception of the cues from the environment and their translation of those cues into an understanding of the situation presented to them by their environmental cues. This understanding is certainly influenced by the lens of personal experience, culture, and education. Each soldier will then take in this series of cues from their five senses and evaluate their situation. The cognitive process the soldier executes to review the presented cues from the environment and determine actions has been built through a lifetime of experience, education and training and is constantly being updated as new life

experiences are added to the soldiers background. The third level of this model proposes that the individual then translates this perception and the developed understanding into a projection of future likely events within their environment. This updated projection of future events impacts the individual's decisions and action choices in the present. This projection relies on the same cognitive interpretation of environmental cues and a developed current perception of the environment as the previous level. The soldier's interpretive lens is based on his prior experience, training and education. This lens facilitates his construction of a likely projected future state.

This work primarily explores a means of representing the perception of the entity. The implemented prototypes also contained simple models for projecting threat movement as described elsewhere (Darken and Anderegg, 2008).

### 2. Related Work

This work builds upon existing models and research results. Current simulations, such as Combat XXI, implement scanning for targets by systematically moving an angular field of view through a larger field of regard. Our reference model replaces this simplistic view of scanning by an attention model as a form of "emergent look around" (Isla and Blumberg, 2002). The form of the threat distribution used to drive the attention model is the "particle filter", which has known operational applications (Ristic et. al., 2004). The assumption of a network of nodes and arcs underlying the terrain representation in a military simulation also has precedents (Reece, 2003).

## 3. Methodology

Data to support the representation of soldier behavior was developed through the execution of a post combat questionnaire and a live virtual experiment.

## 3.1 Post Combat Questionnaire

The post combat questionnaire gathered subject matter expert input to gain insight into the factors essential for modeling close range quick reaction engagements and to address behavioral issues not well represented in current data sets. Questions were designed to illicit responses that

could be used to augment existing data and anticipated data from live virtual experimentation. Questions addressed actions on contact, target acquisition and identification, movement techniques, and combat decision making. The questionnaire was administered to 27 male Army National Guard Soldiers at Camp Shelby, Mississippi. Years of service ranged from 4 to 20 years. Age of the respondents ranged from 21 to 50 years old. The number of direct fire engagements they participated in ranged from 1 to 30. A direct fire engagement was defined as an engagement during which they were either fired upon or fired on an enemy themselves. All respondents participated in direct fire engagements.

The post combat questionnaire results were analyzed using item analysis. The mean and standard deviation were calculated for each item. The hypothesis of marginal homogeneity of the response across the subgroups for each question was examined using the likelihood-ratio chisquare and Pearson chi-square test. These tests are well known and used for the analysis of categorical response and count data and have been used extensively in the analysis of questionnaire and survey data. They differ in their assumption of normality of the estimates. The Pearson chi-square test requires the normality assumption, the likelihood ratio chisquare test does not (Sall, J., Creighton, L., and Lehman, A., 2003). Figure 1 contains a sample question and the analysis results for the role of senses.

Figure 2 shows a representative question and analysis results related to target identification. 56% of the respondents report that target identification takes place at ranges greater than 50 meters. At night this target identification takes place at much closer ranges, with only 4% of respondents reporting identifications at greater than 50 meters. This information can inform the representation of soldier acquisition cycles and be incorporated into the sensing module of the detailed reference model. The results are statistically significant at the .01 level (dark green) and at the .05 level light green.

The question in Figure 3 was designed to inform the movement technique selection associated with soldier behavior in close range quick reaction engagements. The overwhelming majority of respondents reported that they primarily moved using the 3-5 second rush when moving to cover and concealment.

The question in Figure 4 looks to inform the issue of action choice following the identification of a hostile threat. The majority of respondents indicated that they engaged the target first. 38.5% indicated that they informed their teammates first. These responses dominated the other options offered. They are also consistent with the results of similar follow on questions.

The question in Figure 5 seeks to inform the representation of soldier behavior in regard to suppression. Specifically, it seeks to determine

how close the cue, direct fire rounds, needs to be in order to trigger the suppressed behavior. The distribution of responses was broad, but maintained its statistical significance with the greatest number of responses falling in the 6-10m range.

Post combat questionnaires provide a means of soliciting quantifiable subject matter expert input to issues of interest to the analytic and modeling communities. Care should be exercised during their construction to ensure that questions are designed to garner specific pieces of information of use to the modeling effort. These results can be used to fill knowledge and data gaps at relatively low cost.

	Significance of Item Response Chi-Squared Tests		Distribution						
Question	Likelihood Ratio	Pearson	Vision	Hearing	Smell	Touch			
When identifying the presence of a threat (Day), I most often relied on my sense of:	0	0	100.00%	0.00%	0.00%	0.00%			
When identifying the presence of a threat (Night), I most often relied on my sense of:	0.0001	0.0001	44.00%	66.00%	0.00%	0.00%			

Figure 1. Role of senses in threat identification.

	Significance of Item Response Chi-Squared Tests		Distribution						
Question	Likelihood Ratio	Pearson	Less Than 5m	6-15m	16-25m	26-35m	36-50m	Greater Than 50m	
I usually <u>identified (focused on</u> <u>the source of my acquisition)</u> targets at a distance of: Day	0.0077	0.0033	0.00%	0.00%	11.11%	18.52%	14.82%	55.56%	
I usually <u>identified (focused on</u> <u>the source of my acquisition)</u> targets at a distance of: Night	0.0559	0.0721	11.11%	29.63%	7.41%	29.63%	18.52%	3.70%	

Figure 2. Target identification.

	Significance of Chi-Squar	•	Distribution		
Question	Likelihood Ratio	Pearson	3-5 second rush	Low crawl	High crawl
When moving to cover and concealment under fire usually moved by:	0.0001	0.0001	92.00%	4.00%	0.00%

Figure 3. Moving under fire.

Question	Significance of Item Response		Distribution						
	Chi-Squared Tests								
		Pearson		Change					
	Likelihood Ratio			Position					
			Inform	(crouching,		Engage			
			Teammates	kneeling etc)	Seek Cover	Target	Observe		
After identifying a hostile									
target I took the following	0.0001	0.0002	38.46%	3.85%	0.00%	53.85%	3.85%		
action first:									

Figure 4. Action choice on contact.

Question	Significance of Item Response		Distribution							
	Chi-Squared Tests		Distribution							
	Likelihood Ratio	Pearson	1-5m	6-10m	11-15m	16-20m	21-25m	26-30m	>30m	
I felt suppressed when <u>direct</u> <u>fire rounds</u> struck within:	0.0418	0.0486	25.90%	29.60%	18.50%	3.70%	7.40%	3.70%	11.10%	

Figure 5. Suppression from direct fire.

## 3. 2 Live Virtual Experiment

In order to develop additional data on soldier behaviors in close range quick reaction engagement a live virtual experiment was conducted in cooperation with the Soldier Battle Lab at Fort Benning, Georgia. This experiment was conducted their immersive cave facilities. Soldiers were presented with a floor to ceiling frontal screen through which they interacted with the virtual world. The purpose of the experiment was to gain insight into soldier actions on contact, probability of hit, and engagement time over a variety of target ranges and exposure types.

The experiment was executed over five days in March 2007. Ten soldiers from the 29<sup>th</sup> Infantry Regiment participated in the experiment. Their ages ranged from 19 to 35, with 4 having previous combat experience. They were instructed to treat the event as they would a real world situation. They wore a standardized combat ensemble consisting of: interceptor body armor, Kevlar helmet, the Army Combat Uniform, desert boots, six magazines with associated ammo pouches, and canteens. The soldier's received instructions on the use of the M4 surrogate and movement control device embedded in the surrogate to control their movement through the virtual world. Order through the lanes was randomized for each soldier on each day.

Soldiers were instructed to patrol each of the ten lanes and were informed that hostiles could be present. Target cues were presented as soldier's reached predetermined location on each of ten lanes (Figure 6) for each engagement in accordance with the experimental design. If the soldier did not perceive the cue, the cue presentation persisted until the soldier acquired the cue. If the cue was not perceived and the soldier was in danger of by passing or moving beyond the engagement the intensity of the cue was increased.



Figure 6. Top level view of lanes in JCATS terrain database.

The experiment sought to more fully explore the factor space impacting Soldier performance and behavior in close range quick reaction engagements. The factors considered in this experimental design were: Range to target: 0 to 50 meters; Target exposure: 30, 50, and 100 percent exposure of a man sized target; First cue presented: audio or visual cue; Terrain type: interior building, urban canyon, or wooded; Light condition: day or night.

These factors and levels were used to create an experimental design that specified the factor settings for each of ten engagements in ten virtual lanes. Nearly orthogonal Latin hypercube (NOLH) design methodology was used to construct the design matrix. The NOLH design was chosen for use in this effort for a number of reasons. First, it is capable of efficiently sampling from a large number of factors and

levels. Second, the space filling properties of the NOLH are well known, giving it the ability to examine potentially complex response surfaces (Kleijnen, Sanchez, Lucas and Cioppa, 2005).

Data was collected on each soldier throughout each run. This resulted in data on over 4,000 virtual engagements. Data collected included the soldier's location, posture, weapon orientation, shots fired and associated hits, cue type presented and time of presentation.

Our approach to data reduction was to do it entirely via scripts that were run on the raw experimental data each time we built a new variant of the model. The experimental data itself was never altered in any way. The advantage of using scripts is that they provide a clear audit trail from the original data to the data used to construct the model. There were two layers of scripts. First, there was an approximately 500 line Python script that read in the multiple \*.csv raw data files and produced a single \*.csv file containing selected data from the original files augmented with computed variables depending on the original data as described below. Second, there was a short script written in the statistical package R's native scripting language to select independent variables, modestly subset the data, and construct the models. Data where the subject moved more than 900 meters while engaging a single enemy or engaged the enemy from more than 1000 meters were considered outliers and removed. Similarly, engagements with detectionto-trigger times in excess of 10 seconds were neglected.

## 3.3 Analysis Methodology

Data from the experiment was used in conjunction with data from the post combat survey to construct algorithms useful in representing soldier performance and behaviors. These algorithms were developed using logistic regression techniques (see Equation 1). Logistic regression was chosen for a number of reasons: It does not assume homoscedasticity nor linearity of relationship between independent and dependent variables. Additionally, it does not require normally distributed variables (Sall, Creighton, and Lehman, 2005).

$$\ln\left(\frac{p}{(1-p)}\right) = L,$$
(Equation 1)
$$thus \frac{p}{(1-p)} = \exp(L),$$

$$Then, p = \frac{\exp(L)}{[1+\exp(L)]}.$$

Here p is the modeled dependent variable and L is a linear function of the independent variables.

## 4.0 Algorithm Development

The results of the analysis of the experimental data confirmed the results of the post combat questionnaire. Here we define each independent variable used in the models below:

Terrain: Value is one of "Building" (inside a building), "Urban", or "Forest", corresponding to the three conditions used in the live virtual experiments. Time: Value is either "Day" or "Night". The night condition used a coarse approximation to night vision that generally made target detection more difficult as compared to the day condition. Range: The range to the target in meters measured at the moment the subject indicated detection of the target. TargetArea: Value is calculated based on the target range assuming the target is a rectangle of 20" by 69". Cue: Value is either "Visual", meaning that the encounter was arranged so that the subject sees the target first, or "Audible", meaning that the presence of the target is signaled by a gun shot.

Equation 2 describes the relationship between probability of hit and terrain type, day or night condition, exposed target area and the type of cue received first. Engagements conducted in urban terrain at night decreased the probability of hit while a larger presented target area and visual cue presentation first tended to increase the probability of hit.

$$\begin{split} Let \ L &= -1.471 - 2.109e - 01*(Terrain = Urban) \\ &- 1.787e - 01*(Time = Night) \\ &+ 3.826e - 06*(Target\ Area) \\ &+ 1.066e - 01*(Cue = Visual), \end{split}$$
 (Equation 2) 
$$Then, \ p_{hit} = \frac{\exp(L)}{[1 + \exp(L)]}.$$

Equation 3 describes the relationship between the time from target detection to engagement and terrain type, day or night condition, range to the target and the type of cue received first. The time to complete the engagement following a successful target detection increased as range to the target increased in both forested and urban terrain conditions, but was surprisingly decreased by a nighttime condition.

$$\label{eq:local_$$

Equation 4 shows the relationship between the action choice to move first and terrain type, day or night condition and range to the target. The probability of moving first decreases as the range to the target decreases and is positively impacted by a nighttime condition.

$$\begin{aligned} Let \ L &= 3.383447 + 1.617288*(Terrain = Forest) \\ &+ 1.626717*(Terrain = Urban) \\ &+ 0.828505*(Time = Night) \\ &- 0.011973*Range, \end{aligned}$$
 (Equation 4) 
$$\begin{aligned} Then, \ p_{move} &= \frac{\exp(L)}{[1 + \exp(L)]}. \end{aligned}$$

If the choice is made to not move, the following model separates the remaining cases into those where the firing is done from a kneeling position and those where firing is executed from a crouching position. Equation 5 shows that the probability of kneeling decreases at night but does increase as range to the target increases.

$$\begin{aligned} & Let \ L = -6.91101 + 2.13164*(Terrain = Forest) \\ & -0.93547*(Time = Night) \\ & +0.0378*Range, \\ & Then, \ p_{kneel} = \frac{\exp(L)}{[1 + \exp(L)]}. \end{aligned}$$
 (Equation 5)

These algorithms, developed from live virtual experimentation, became the basis for the action selection component of the reference model.

#### 5. Soldier Reference Model

The objective of this portion of the research is to develop a highly detailed reference model of soldier behaviors and actions that can be used as a vehicle to transfer developed algorithms and knowledge into current large scale simulation models. This reference model (see Figure 7) can also serve to aid in the identification of gaps in the knowledge and data required to represent soldiers in models and simulation. The focus of

this model is on visual attention (gaze control), situation awareness, target detection, and initial reaction on contact. The model as a whole is conceptual, though large pieces, namely the action on contact models above and the decision-theoretic attention model described below, were implemented as software prototypes.

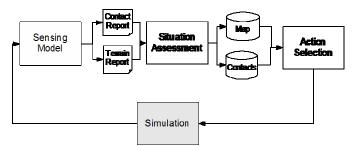


Figure 7 Conceptual reference model.

# 5.1 Sensing Model

The sensing model is responsible for the generation of contact and terrain reports. Terrain reports are simplest. When a new part of the map is viewed for the first time, the terrain reports corresponding to that part of the map are immediately generated. A terrain report consists of a list of discrete locations (sometimes known as waypoints or navigation nodes) that have come into view of the model. Knowledge of "contacts", i.e. other simulation entities, requires more computation. When the sensing model detects a state of affairs that could cause one entity to be aware of another, such as one entity being within another's sensing range (e.g. field of view for visual sensing, or auditory range for footsteps or gun shots, if represented), a target acquisition model is applied to determine whether and when the agent perceives anything. We have previously described our approach to target acquisition of the context of simulations having a detailed 3D representation of the environment (Darken and Jones, 2007).

Contact reports are the sole information source input to the model regarding the positions of moving entities. A minimal contact report consists of a contact ID (arbitrary identifier used to distinguish a particular contact), a time stamp, and a position distribution corresponding to this particular sensing of the contact. Contact models may be augmented in the future with additional features describing what was sensed in order to discriminate contact types. We do not handle

contact aliasing in this model (i.e. ambiguous situations where a sensation might be from one of several different contacts).

#### 5.2 Situation Assessment

The situation assessment module is the part of the model that updates the contact and map representations based on contact and terrain reports and updates the entities situation awareness. When a terrain report is received, the module flags those navigation nodes contained in the report as known to the soldier model. This corresponds to updating the entities perception of its environment, level 1 situation awareness (Endsley, M. R., 2001). Those nodes might now be used for route planning and for cover taking. This component also represents the complete situation awareness of the model with regard to other simulation entities. While knowledge of the location of friendly and neutral entities can potentially be handled by the same model, we focus on hostile entities only in this treatment. The contact representation is a probability distribution representing the contact's position. For this version of the model, there is one distribution for each known contact.

When a contact report is received, if it describes a completely new contact, the position distribution in the contact report is sampled to populate the model's position distribution (labeled "Contacts" in the architecture diagram of Figure 7) with a set of representative position hypotheses and their relative likelihoods. This representation of the contact constitutes a probability distribution of the contact's position known as a "particle filter" (Darken and Anderegg, 2007). If the contact report describes a contact that is already known, the model's position distribution is updated to incorporate the new information. The contact report is treated as an independent of previous reports, and the position distribution is updated.

Let X be a random variable describing the contact position at some point in time. The information in all previous contact reports that was used to construct the distribution of X we encapsulate as random variable A. Let random variable B represent the new information in the contact report just received. We assume that B is independent of A. We want to update the contact position to be P(X|A,B), i.e. the position conditioned on the information in both A and B.

Applying Bayes' rule to P(X|A,B) yields

$$P(X \mid A, B) = P(A, B \mid X)P(X)/P(A, B)$$
 (Equation 6)

But since A and B are independent,

$$P(X \mid A, B) = P(A \mid X)P(B \mid X)P(X)/P(A, B)$$
 (Equation 7)

Applying Bayes' rule to P(A|X) and P(B|X) and again noting the independence of A and B yields,

$$P(X \mid A, B) = P(A \mid X)P(B \mid X)/P(X)$$
 (Equation 8)

Let L be the finite set of positions which are the range of X. If we lack knowledge of P(X), the prior distribution over X without A or B, it is convenient to let P(X) be the uniform distribution, i.e. a constant. The proper value of the constant is obvious since P(X|A,B) must sum to one, and so the update for the contact position assuming a uniform prior is:

$$P(X \mid A, B) = P(X \mid A)P(X \mid B) / \sum_{x \in I} P(x \mid A)P(x \mid B)$$
 (Equation 9)

When no contact report is received, but potential threat positions are in view, the probability that the threat is actually at one of those positions is decreased. To compute the amount of decrease, we assume that the model has an independent chance to detect a target each time it is brought into view (assuming the model has looked away in the meantime). Let the initial probability that there is a threat at a given position be  $p\theta$ . Assume that our target detection model gives P(t) as the probability that this target will be detected in no more than t seconds. If we observe this position for s seconds without receiving a contact report indicating a target detection, the remaining probability that there is a threat there is p0(1-P(s)).

The map representation consists of a single bit per navigation node, representing whether the model has knowledge of the node or not. Knowledge of a navigation node is either zero or complete. The model assumes that a special representation of the terrain is available to support the model, specifically a navigation graph. A navigation graph consists of a set of nodes (often called navigation nodes or waypoints) which is a subset of positions where a model may position itself. Each node is aware of any sufficiently nearby nodes in the graph that the model could move to in a straight line. The navigation graph can save computation at runtime by making it possible to quickly plan paths between arbitrary nodes that avoid static

obstacles in the synthetic environment, such as walls, boulders, trees, etc. Additionally, we assume that each node also contains six values representing the amount of cover provided from a set of 60 degree arcs covering all directions. The navigation graph and cover data can be constructed by an automated process as previously described (Darken, Anderegg and McDowell, 2007) and (Darken, 2007).

#### 5.3 Action Selection

The action selection module controls the gaze direction of the model using a decision theoretic model. It also makes the decision of how to respond to quick reaction engagements using the statistical models described above. We take a decision theoretic approach to gaze control, choosing the direction of maximum threat density in which to gaze. Threat density is computed by aggregating by angular sector the probability of all threat positions in the contact model that could be brought into view by adjusting gaze (i.e. threat positions that are behind an obstacle and would require walking to a new position to bring them into view are ignored). The center of the sector containing the largest threat density is chosen as the new gaze direction. When a target is acquired, this module uses the logistic regression models described above to choose

#### 6. Conclusions and Future Work

This paper presented our model of soldier behavior up to his initial choice of action in a close range encounter. While the model as a whole is conceptual and intended to drive future investigations, substantial chunks of it have been implemented to the point where we believe that a proof in principle has been provided. We expect that future work will allow us to improve the components of this detailed model and to cast it into simpler forms as required for inclusion in current simulations concerned with soldier behavior.

### 7. References

Darken, C. (2007). Level Annotation and Test by Autonomous Exploration. *Proceedings of Artificial Intelligence and Interactive Digital Entertainment (AIIDE)* 2007.

Darken, C., Anderegg, B. and McDowell, P. (2007). Game AI in Delta3D. *Proceedings of Computational Intelligence in Games (CIG)* 2007.

Darken, C. and Anderegg, B. (2008). Particle Filters and Simulacra for More Realistic Opponent Tracking. In S. Rabin (Ed.), *Game AI Programming Wisdom 4*, Boston: Charles River Media.

Darken, C. and Jones, B. (2007). Computer Graphics-Based Target Detection for Synthetic Soldiers. *Proceedings of Behavior Representation in Modeling and Simulation* (BRIMS) 2007.

Endsley, M. R. (2001). Designing for situation awareness in complex systems. *Proceedings of the Second international workshop on symbiosis of humans, artifacts and environment, Kyoto, Japan.* 

Isla, D. and Blumberg, B. (2002). Object persistence for synthetic creatures. *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*.

Kleijnen, J.P.C., Sanchez, S., Lucas, T., and Cioppa, T. (2005). A user's guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing*, 17(3), 263-289.

Reece, D. (2003). Movement behavior for Soldier Agents on a Virtual Battlefield. *Presence*, 12(4), 387-410.

Ristic, B., Arulampalam, S. and Gordon, N. (2004). *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Norwood, Massachusetts: Artec House Radar Library.

Sall, J., Creighton, L., and Lehman, A. (2003). *JMP Start Statistics*. Belmont, California: Brooks/Cole—Thomson Learning.. **Author Biographies** 

MAJ JONATHAN K. ALT is an Operations Analyst assigned to TRADOC Analysis Center – Monterey. He received an M.S. in Operations Analysis from the Naval Postgraduate School in 2006, holds an M.Ed. from the University of Georgia and a B.S. in Engineering Management from the United States Military Academy. Prior to becoming an Operations Analyst he served as an infantry officer in the 3<sup>rd</sup> Infantry Division and the 101<sup>st</sup> Airborne Division.

# CHRISTIAN J. DARKEN, PhD, is an

Associate Professor of Computer Science at the Naval Postgraduate School, where he also collaborates intensively with the MOVES Institute. Previously he was Project Manager of the Decision Support Systems project and Senior Member of Technical Staff at Siemens Corporate Research in Princeton, NJ, where he was variously associated with the Learning Systems, Adaptive Information and Signal Processing, and Software Engineering Departments. He was also a programmer on the team that built Meridian 59, the first 3D massively-multiplayer game. He received his Ph.D. in Electrical Engineering from Yale University in 1993, and previously received the M.S. and M. Phil. in Physics from the same institution.