The Extended Continuous Evaluation Reliability Growth Model

Larry H. Crow, Ph.D., Crow Reliability Resources

Key Word: Continuous Evaluation, Operational Testing, Reliability Management, Reliability Growth, Mission Profiles

SUMMARY & CONCLUSIONS

In 2007 the Defense Science Board convened a Task Force to investigate and make recommendations on changes that would increase the reliability of military systems. In 2008 a report was issued by the Task Force which gives the Task Force findings and recommendations. The Task Force found that the "lack of continuous RAM improvement during design, and the resulting low initial MTBF and low Growth Potential are the most significant reasons that systems are failing to meet their operational suitability requirements." In addition, the Task Force recommended that "RAM, to include a robust reliability growth program, be a mandatory contractual requirement and document progress as part of every major program review." The Task Force also noted that, "More importantly, an operational perspective earlier in the developmental process has often proven to be a catalyst to early identification and correction of problems." See Ref. 3.

As a result of these finding it is clear that reliability growth tests need to be implemented and conducted closer to operational, real-word conditions in order to evaluate and grow the reliability under conditions close to actual use conditions. This is referred to as "Operational-Like" reliability growth test and represents a new direction within the Department of Defense. A structured Operational-Like reliability growth test would be testing conducted in accordance with the Mission Profile Testing methodology presented in Crow Ref. 2. The author, as a member of this Task Force, concurs with these findings and recommendations. This paper presents a model that is design for continuous reliability growth evaluation over a single or multiple test phases. The model has applications to reliability growth testing in general and is particularly appropriate for Operational-Like testing.

1. INTRODUCTION

The model discussed in this paper is an extension of the basic Extended reliability growth model presented by Crow in Ref. 1. The application of the model is general but it is particularly appropriate for the mission profile reliability growth testing discussed in Crow Ref. 2. The reader is referred to both of these papers for additional background. The main objectives of the model discussed in this paper are to be able to continuously evaluate and manage the reliability growth of a system across multiple test phases and to accommodate failures that are likely to be seen during Operational-Like testing.

Under Operational-Like mission profile testing there are important differences from a typical reliability growth development test. A key difference is that under a typical reliability growth development test the failure definition is often restricted to hardware or software failures due to design issues. The main focus of the reliability growth development test is to increase the reliability of the hardware and software design of the system. Under Operational-Like mission profile reliability growth testing this definition and the focus of the testing is broader. Under Operational-Like reliability growth testing the focus is not only failures due to hardware or software design causes but also failures due to human factor causes. The reason the human factor causes is added to the failure definition for Operational-Like testing is that all these failures are what the user sees, and for many programs under development the percent of the total system failures due to human factor causes ranges between 30% and 50 %. This percent is significant, and implies that in order to attain the operational reliability requirement the human factors failure causes must be addressed early by corrective action. This is often not the case under a typical reliability growth program. However, under an Operational-Like reliability growth testing program these failure causes are addressed and counted in the reliability assessment and in an estimate of the growth potential. In particular, failures under Operational-Like reliability growth testing would typically include failures due to (1) Hardware Design (2) Software Design (3) Manufacturing (4)Vendor Quality (5) Operator Error (6) Operator Maintenance.

Under the Operational-Like mission profile reliability growth testing structure discussed in Ref. 2, corrective actions would be expected to be incorporated into the system at several predetermined times during the testing. This means that the testing is conducted in phases, but the phases are all related. Many reliability growth testing programs are conducted over several test phases. We therefore want a model that will utilize all the data across test phases to best use the entire data set for valid assessments. Current reliability growth models are, in general, designed for application during a single test phase where all the corrective actions are incorporated during the test phase, at the end of the test phase, or both. There generally is no provision for carrying corrective actions across test phases. The model discussed in this paper has key features that address these important issues.

Notation

λ Scale parameter for Crow (AMSAA) model

978-1-4244-5103-6/10/$26.00 ©2010 IEEE

β	Shape parameter for Crow (AMSAA) model
T	End of a test phase
MTBF	Mean time between failures
d_i^{Nom}	The Nominal Effectiveness Factor for ith mode
d_i^{Act}	The Actual Effectiveness Factor for ith mode
$MTBF_D$	Demonstrated MTBF
Xj	The j-th successive failure time
N	Total number of failures

2. BACKGROUND

For the basic Extended model, Ref.1, there are Type A modes that if seen are not corrected, Type BC modes that if seen are the type of failure mode that will always be corrected during test, and Type BD modes that if seen the corrective action is delayed and will always be corrected after all the testing is completed at time T. Testing is not continued with this model past time T. Type BC modes are either fixed at the time of failure or at some later time during the test, but before time T. For the basic Extended Model it is required that all BD failure modes be corrected when the testing is stopped at time T.

With Operational-Like testing the Type BC and Type BD definitions above are not particularly suited for data and corrective actions over multiple test phases. This is because with multiple test phases a mode that was once a BD mode may later be reclassified as a BC mode. This is, however, inconsistent with the basic Extended model definitions. Therefore, in the Extended Continuous Evaluation model the definitions of Type BC and Type BD modes are modified.

For the Extended Continuous Evaluation model the definition of Type BC models changes: A Type BC failure mode is a failure mode that receives a corrective action at the time of failure and before the testing starts up again. It would be expected that a Type BC mode corrective action would result in reliability growth during the test. A Type BC failure mode generally requires no extensive root cause failure analysis and therefore can be corrected quickly. Type BC modes are generally easy to fix and are usually quality, manufacturing, operator, etc. type failure modes. For the Extended Continuous Evaluation model the definition of Type BD modes also changes: A Type BD failure mode is a failure mode that receives a corrective action at a time after the first occurrence of the failure mode. Type BD failure modes typically require failure analysis, and time to fabricate the corrective action. During the period (0, T) into the test there may be BD modes with corrective actions incorporated into the systems and some BD modes seen but not yet fixed. The definition of Type A failure modes remains the same, and are failure modes such that if seen will not receive a corrective action. The Type A, Type BC and Type BD failure modes, together with the effectiveness factors, define the management strategy and determines the system growth potential.

It is, also, noted that many human factor failures are corrected at the time of failure. Typically the problem must be corrected in order to continue the testing. The root cause of a human factor failure may also be due to a design issue and is corrected later. For Type BD hardware failures the failed item is replaced at the time of failure, and a corrective action is incorporated later. Software corrective actions are generally after the time of failure.

3. THE EXTENDED CONTINUOUS EVALUATION MODEL

The new definitions of Type BC and Type BD failure modes aligns the Extended Continuous Evaluation model to operational type failures where the BC modes will typically be due to human factor causes, and the Type BD failure modes are typically due to hardware design and software design causes. This places more clarity and focus on the BC and BD failure modes for analysis and management. If human factor failures are not counted in the failure definition then most failures may be expected to be typically classified as BD failure modes. The other key area that needs to be addressed is analyzing data across test phases. This is the purpose of an additional parameter in the model.

At the end of the test phase the basic Extended model requires that all remaining Type BD modes be corrected as delayed corrected actions. This occurs with probability one. This is not required with the model presented in this paper. During a test phase ending at time T the Extended Continuous Evaluation model allows for delayed corrected actions to be fixed before time T, at time T, or, in particular, after time T during a later test phase. The definition of "delayed" is expanded to include all Type BD failure modes corrected after the time of failure but not necessarily at time T or before. Under this definition of a BD mode it follows that whether or not a BD mode seen during (0, T) is corrected after time T depends on many factors and can be considered as random. Under the Extended Continuous Evaluation model the parameter p(T) is the probability that a BD failure mode seen during (0, T) will be corrected after time T. The probability 1- p(T) is the probability that a BD failure mode seen during (0, T) will be corrected before time T. A BD mode seen during (0, T) may be either corrected or not corrected at time T according to the probability p(T).

For nomenclature purposes we denote BD failure modes that have been corrected by time T as BDC (delayed but now corrected) and those BD modes that have not received a corrective action by the stop of testing at time T as BDD (delayed but not corrected). As noted above, the parameter p(T) is the probability that a BD mode seen during (0, T) is a BDD mode at time T. The probability that a BD mode seen during (0, T) is a BDC mode at time T is 1- p(T).

With the Extended Continuous Evaluation model reliability analyses may be conducted multiple times during the test program and, in particular, at the end of each test phase. A test phase usually ends at a planned stopping point in the test program. There are generally delayed corrected actions incorporated into the system before the next test phase starts. An additional general feature of a reliability growth test phase is that there are usually interim requirements or goals to be meet. Therefore, there are generally assessments of the current reliability and a projection of the reliability expected in the next test phase.

Each time an analysis is conducted the Extended Continuous Evaluation model allows for several key metrics to be calculated, including growth potential, that support the intent of the Defense Science Board Task Force recommendations. At the end of the j-th test phase at time T_j we can conduct all of these analyses and consequently better manage the program.

The Extended Continuous Evaluation model and metrics are illustrated by an example. With a test program consisting of several test phases these metrics would be calculated multiple times. The example illustrated below could be at the end of any test phase.

4. *EXAMPLE ILLUSTRATING THE EXTENDED CONTINUOUS EVALUATION MODEL*

Suppose the reliability growth testing is specified for 3500 hours and at time T = 400 hours into the testing, the testing is stopped in order to update the system configuration with corrective actions. That is, the end of a test phase. Between time 0 and time T = 400 corrective actions were incorporated for some of the failure modes and additional corrective actions are planned to be incorporated at time T = 400. With the Extended Continuous Evaluation model assessments can be made throughout the testing. After the testing is resumed at time 400, additional assessments can be made and at time T = 3500 all remaining BD modes must all be corrected.

At an assessment point, for example T = 400, the corrective action status of all failure modes seen to-date is updated, as in Table 1. A failure mode corrected at the time of failure is denoted by a BC mode. Any failure mode that is not planned to be corrected during the 3500 hour test is denoted by an A mode. All failure modes that are planned to be corrected after the time of first occurrence are denoted as a Type BD mode. Among the BD modes those that have been corrected by the stop of the testing at time T = 400 are further denoted as a BDC mode. Those BD modes that have not been corrected by the stop of the testing at time T = 400 are denoted as a BDD mode. Of these BDD modes we can now chose to incorporate corrective actions for some before testing resumes, and defer corrective actions for the others to a later time between T = 400 and up to and including T = 3500. At time T = 400 we wish to get an updated assessment of our progress and the impact of the proposed corrective actions.

In Table 1, there are a total of 17 distinct Type BD modes. Of these 17 failure modes 5 have been corrected (Type BDC modes). The BDC modes are 1, 9, 10, 15 and 16. There are 12 Type BD modes not yet fixed (Type BDD modes). These are modes 2, 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 17. After the testing has stopped at time T = 400 we will incorporate corrective actions for BDD modes 2, 4, 6, 11, 14, 17. After these corrective actions are incorporated and testing resumed (T > 400) each of these BDD modes will be denoted as a BDC mode, that is, delayed and now corrected. Type BD modes 3, 5, 7, 8, 12, 13 are still denoted as BDD modes until corrected.

Time to Failure	Failure Mode Status	Failure Mode Name	Time to Failure	Failure Mode Status	Failure Mode Name
0.7	BDC	1	192.7	BDD	12
15	BDD	2	213	A	35
17.6	BC	23	244.8	A	36
25.3	BDD	3	249	BDD	13
47.5	BDD	4	250.8	A	35
54	BDD	5	260.1	BDD	2
54.5	BC	24	273.1	A	35
56.4	BDD	6	274.7	BDD	6
63.6	A	34	282.8	BC	32
72.2	BDD	5	285	BDD	14
99.2	BC	25	315.4	BDD	4
99.6	BDD	7	317.1	A	34
100.3	BDD	8	320.6	A	36
102.5	A	34	324.5	BDD	12
112	BDC	9	324.9	BDC	10
112.2	BC	26	342	BDD	5
120.9	BDD	2	350.2	BDD	3
121.9	BC	27	355.2	BC	33
125.5	BDC	10	364.6	BDC	10
133.4	BDD	11	364.9	A	35
151	BC	28	366.3	BDD	2
163	BC	29	379.4	BDC	15
174.5	BC	30	389	BDC	16
177.4	BDC	10	394.9	A	36
191.6	BC	31	395.2	BDD	17

Table 1. Failure Times in First 400 Hours of Test, Failure Mode Status, and Failure Mode Name

In Table 2 we assigned the Nominal Fix Effectiveness Factors for all the 12 Type BDD modes. A Nominal Fix Effectiveness Factor is the expected fraction decrease in the failure mode failure rate after a corrective action is implemented. There is also an assigned Actual Fix Effectiveness Factor which is equal to the Nominal Fix Effectiveness Factor if the corrective action is implemented at time T = 400 and is equal to zero if the corrective action is not implemented at time T = 400. Let d_N and d_A denote the average of the nominal and actual effectiveness factors. In the example

$$d_N = 0.699 \tag{1}$$

and

$$d_A = 0.346 . \tag{2}$$

Each time we make an assessment with the Extended Continuous Evaluation model we can calculate each of these metrics:

a) Current Demonstrated MTBF
b) Nominal Growth Potential
c) Nominal Average EF
d) Nominal Projection if BDD modes are corrected with Nominal EFs
e) Actual Growth Potential
f) Actual Average EF
g) Actual Projection if BDD modes are corrected with Actual EFs

h) Rate of discovery.

Time to First Failure	Failure Mode Status	Failure Mode Name	Nominal EF	Actual EF
15	BDD	2	0.67	0.67
25.3	BDD	3	0.72	0
47.5	BDD	4	0.77	0.77
54	BDD	5	0.77	0
56.4	BDD	6	0.87	0.87
99.6	BDD	7	0.92	0
100.3	BDD	8	0.5	0
133.4	BDD	11	0.74	0.74
192.7	BDD	12	0.70	0
249	BDD	13	0.63	0
285	BDD	14	0.64	0.64
395.2	BDD	17	0.46	0.46

Table 2. Failure Times and Nominal and Actual Effectiveness Factors for the Distinct BDD Failure Modes

The current demonstrated MTBF addresses the question: What is the reliability that is currently being demonstrated at time T = 400? This is given by the Crow (AMSSA) model. (See Ref. 1.) In Table 1 there are N = 50 total failures. Let X_i denote the i-th failure time in Table 1, i=1,…,50. Then, (using the unbiased estimate of beta) the Crow (AMSAA) model failure intensity at T = 400 hours is

$$\hat{\lambda}_D = \hat{\lambda}\hat{\beta}T^{\hat{\beta}-1} \tag{3}$$

where

$$\hat{\beta} = \frac{N-1}{\sum_{J=1}^{N}\ln\left(\frac{T}{X_i}\right)}, \tag{4}$$

and

$$\hat{\lambda} = \frac{N}{T^{\hat{\beta}}}. \tag{5}$$

In this example,

$$\hat{\beta} = .9669, \tag{6}$$

$$\hat{\lambda} = .1524 \tag{7}$$

and the demonstrated failure intensity is

$$\hat{\lambda}_D = 0.1209. \tag{8}$$

The demonstrated, or achieved MTBF is

$$MTBF_D = \frac{1}{\hat{\lambda}_D} = 8.27. \tag{9}$$

If we put corrective actions into the system for some of the BDD modes what is our estimate of the new reliability? This is a projection. The Nominal Projection is the estimate if we fix all of the BDD modes at time 400, and the Actual projection is the estimate if we only fix those that have designated.

We first calculate the Nominal Growth Potential Factor

$$\lambda_{NGPFactor} = \sum_{i=1}^{K}(1-d_i^{Nom})\frac{N_i}{T} \tag{10}$$

where K is the total number of distinct BDD modes at time T = 400, d_i^{Nom} is the assigned Nominal EF for the i-th BDD mode at time T =400, and N_i is the total number of failures over (0, 400) for distinct BDD mode i. In the example K = 12, and the EFs are given in Table 2.

In the example,

$$\lambda_{NGPFactor} = \sum_{i=1}^{K}(1-d_i^{Nom})\frac{N_i}{T} = 0.015000. \tag{11}$$

Next, we calculate the estimate of p(400). This estimate is the total number of distinct BDD modes at time T = 400 divided by the total number of distinct BDC modes plus the total number of distinct BDD modes. In this example p(400) = (12/17) = 0.7059

The BDD mode failure intensity at time T= 400 is λ_{BDD} = total number of BDD failures (classification at the end of test at T = 400) divided by time 400. In the example

$$\lambda_{BDD} = 21/400 \tag{12}$$

or

$$\lambda_{BDD} = 0.0525. \tag{13}$$

The discovery function $h(t) = \lambda\beta t^{\beta-1}$ is calculated using all the first occurrences of the 17 BD modes. This is the rate in which new, distinct, Type BD modes are being discovered during the test. The M = 17 first failure times, Z_i, corresponding to discovering a new, distinct, Type BD mode (both BDC and BDD modes) are .7, 15, 25.3, 47.5, 54, 56.4, 99.6, 100.3, 112.0, 125.5, 133.4, 192.7, 249.0, 285.0, 379.4, 389.0, 395.2.

The unbiased estimate of beta for the h(t) function is:

$$\beta^* = \frac{M-1}{\sum_{i=1}^{M}\ln\left(\frac{T}{Z_i}\right)} = \frac{16}{\sum_{i=1}^{17}\ln\left(\frac{400}{Z_i}\right)} = 0.6055. \tag{14}$$

Beta must be less than one for reliability growth. The h(t) function at time T = 400 is

$$h(400) = \beta^*\frac{M}{T} = 0.0257. \tag{15}$$

The classification of the A, BC and BD modes and the effectiveness factors define the management strategy. The growth potential is the maximum reliability that can be attained with the current management strategy. The nominal growth potential is the growth potential if all BD modes in the system were seen and corrected with the nominal effectiveness factors. The actual growth potential is the growth potential if all BD modes in the system were seen and corrected with the actual effectiveness factor strategy that has been demonstrated up to time T.

The Nominal Growth Potential failure intensity is

$$\lambda_{NGP} = \lambda_D - \lambda_{BDD} + \lambda_{NGPFactor} + d_N \cdot p \cdot h(T) - d_N \cdot h(T) \tag{16}$$

For this example,

$$\lambda_{NGP} = 0.0781 \tag{17}$$

The Nominal Growth Potential MTBF is

$$M_{NGP} = \frac{1}{\lambda_{NGP}}. \tag{18}$$

For this example,

$$M_{NGP} = 12.8. \tag{19}$$

The Nominal Growth Potential MTBF is an estimate of the maximum reliability that is attainable for the system with the current design and management strategy. The overall system MTBF requirement or goal must be below the Nominal Growth Potential MTBF.

The Nominal Projection estimates the failure intensity and MTBF if all seen BDD modes are corrected at time T. The Nominal Projected Failure Intensity at time T is

$$\lambda_{NP} = \lambda_{NGP} + d_N \cdot h(T) \quad (20)$$

In the example

$$\lambda_{NP} = 0.0961. \quad (21)$$

The Nominal Projected MTBF at time T is

$$MTBF_{NP} = \frac{1}{\lambda_{NP}}. \quad (22)$$

In the example the Nominal Projected MTBF is

$$MTBF_{NP} = 10.4. \quad (23)$$

If all 12 BDD modes 2, 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 17 were fixed at time T = 400 then the estimate of the new MTBF is 10.4, given by the Nominal Projected MTBF. However, we are only going to fix BDD modes 2, 4, 6, 11, 14 and 17 at time T = 400. The Actual Projected MTBF is the estimate of the new MTBF after we fix only modes 2, 4, 6, 11, 14 and 17 at time T = 400.

The Actual Growth Potential Factor is

$$\lambda_{AGPFactor} = \sum_{i=1}^{K} (1 - d_i^{Act}) \frac{N_i}{T} \quad (24)$$

where K is the total number of distinct BDD modes at time T, d_i^{Act} is the assigned actual EF for the i-th BDD mode at time T and N_i is the total number of failures over (0, T_j) for distinct BDD mode i. In the example K = 12, and the actual EFs are in Table 2. In the example,

$$\lambda_{AGPFactor} = \sum_{i=1}^{K} (1 - d_i^{Act}) \frac{N_i}{T} = .03300. \quad (25)$$

The Actual Growth Potential Failure Intensity is

$$\lambda_{AGP} = \lambda_D - \lambda_{BDD} + \lambda_{AGPFactor} + d_A \cdot p \cdot h(T) - d_A \cdot h(T) \quad (26)$$

For this example,

$$\lambda_{AGP} = 0.0987. \quad (27)$$

The Actual Growth Potential MTBF is

$$M_{AGP} = \frac{1}{\lambda_{AGP}}. \quad (28)$$

For this example,

$$M_{AGP} = 10.13. \quad (29)$$

The Actual Growth Potential MTBF is an estimate of the maximum reliability attainable for the system with the current design and the management strategy that is demonstrated by the actual effectiveness factors.

The Actual Projected Failure Intensity at time T is

$$\lambda_{AP} = \lambda_{AGP} + d_A \cdot h(T) \quad (30)$$

In the example,

$$\lambda_{AP} = 0.1076. \quad (31)$$

The Actual projected MTBF at time T is

$$MTBF_{AP} = \frac{1}{\lambda_{AP}}. \quad (32)$$

In the example, the Actual projected MTBF

$$MTBF_{AP} = 9.29. \quad (33)$$

Graphically, these results are given in Figure 1 below.

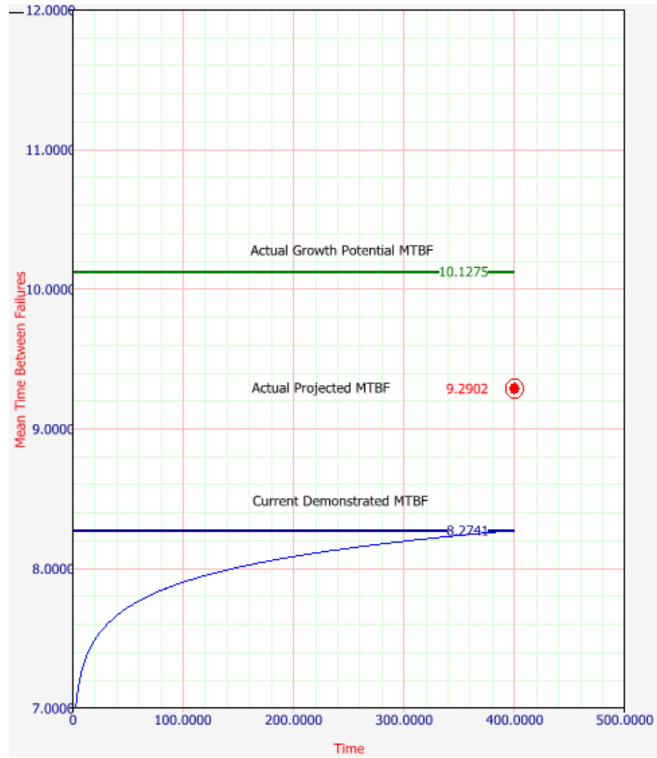

Figure 1. The Demonstrated MTBF , Actual Projected MTBF and Actual Growth Potential MTBF at T =400

These calculations can be updated continuously throughout the entire reliability growth test and across the test phases. Also, the model can be implemented using grouped data over the test phases. At the completion of the reliability growth test all remaining BDD modes would be expected to be corrected.

REFERENCES

1. L. H. Crow, An Extended Reliability Growth Model for Managing and Assessing Corrective Actions. Proceedings of the 2004 Annual RAM Symposium, January 2004, Los Angeles, CA.
2. L. H. Crow, A Methodology for Managing Reliability Growth during Operational Mission Profile Testing,. Proceedings of the 2008 Annual RAM Symposium, January 2008, Las Vegas, NV.
3. Defense Science Board, Report of the Defense Science Board Task Force on Developmental Test and Evaluation, May 2008.

BIOGRAPHY

Larry H. Crow, Ph. D.
109 Clifts Cove Blvd.
Madison, AL 35758 U. S. A.

Internet (e-mail) Crowrel@knology.net

Dr. Larry H. Crow is president of CRR. Previously Dr. Crow was VP, Reliability & Sustainment Programs, at ALION Science and Technology, Huntsville, AL. From

1985 to 2000, Dr. Crow was Director, Reliability, at General Dynamics ATS -- formally Bell Labs ATS. From 1971-1985, Dr. Crow was chief of the Reliability Methodology Office at the US Army Materiel Systems Analysis Activity (AMSAA). He developed the Crow (AMSAA) reliability growth model, which has been incorporated into US DoD handbooks, and national & international standards. He chaired the committee to develop Mil-Hdbk-189, 1981, *Reliability Growth Management,* and is the principal author of that document. Dr. Crow is a Fellow of the American Statistical Association, and the Institute of Environmental Sciences and Technology. He is a Florida State University Alumni Association Distinguished Alumnus and the recipient of the FSU "Grad Made Good" Award for the Year 2000, the highest honor given to a graduate by Florida State University.