EFFECTS OF SESSION DURATION ON FUNCTIONAL ANALYSIS OUTCOMES

MICHELE D. WALLACE AND BRIAN A. IWATA

THE UNIVERSITY OF FLORIDA

We examined the extent to which variations in session duration affected the outcomes of functional analyses. Forty-six individuals, all diagnosed with mental retardation and referred for assessment and treatment of self-injurious or aggressive behavior, participated in functional analyses, consisting of repeated exposure to multiple test conditions during 15-min sessions. For each set of assessment data, new data sets based on session durations of 10 and 5 min were prepared by deleting data from the last 5 and 10 min, respectively, of each session. Each graph ($N = 138$) was then reviewed individually by graduate students who had previous experience conducting and interpreting functional analyses, but who were blind to both participant identity and session duration. Interpretations of behavioral function based on the 10- and 5-min data sets were then compared with those based on the 15-min data sets. All of the 10-min data sets yielded interpretations identical to those based on 15-min data sets. Interpretations based on the 5-min and 15-min data sets yielded three discrepancies, all of which were the result of increased response rates toward the latter parts of sessions. These results suggest that the efficiency of assessment might be improved with little or no loss in clarity by simply reducing the duration of assessment sessions.

DESCRIPTORS: assessment, functional analysis, observation procedures

---

Functional analysis methodology has become a common research and assessment tool for identifying environmental variables that maintain a wide range of behavior disorders. In a typical assessment (e.g., Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/ 1994), rates of behavior observed under one or more test conditions, in which suspected maintaining variables are present, are compared with those observed under a control condition, in which maintaining variables are absent. These observations are often conducted during 10- or 15-min sessions, which are repeated until response differentiation is evident. Despite the demonstrated utility of this approach to assessment, it has been characterized by some as being either too complex or too time consuming (Durand & Crimmins, 1988; Lennox & Miltenberger, 1989; Pyles, Riordan, & Bailey, 1997; Repp, Singh, Olinger, & Olson, 1990). For example, in a recent study in which data were summarized for over 150 functional analyses (Iwata et al., 1994), it was noted that the mean length of assessment was 26 sessions, which represented approximately 6.5 hr of observation.

Several attempts have been made to increase the efficiency with which functional analyses can be conducted through procedural variations of one sort or another. The most notable of these is often referred to as the "brief functional analysis," first reported by Northup et al. (1991). These authors demonstrated the use of single exposures to test and control conditions during sessions lasting for 5 to 10 min, such that the entire assessment could be completed during a 90- min outpatient evaluation. Results obtained for 2 of the participants in that study sug-

gested multiple functions for their aggressive behavior, and it is possible that clearer results might have been obtained by collecting additional assessment data. Because the 3rd participant was exposed to only one test condition, the possibility that behavior was maintained by other sources of reinforcement could not be ruled out. Thus, although the study made a significant contribution by illustrating several variations in functional analysis methodology, the utility of those variations was not clearly demonstrated. Subsequently, Derby et al. (1992) presented summary data for 79 brief functional analyses conducted in an outpatient setting, in which assessment was based on test and control conditions conducted once each for 10 min, followed by brief contingency reversals. The authors were able to identify the function of behavior problems in 39 of the cases (49%). However, because results from these brief analyses were not compared with a more definitive standard, the extent to which more complete analyses would have produced clearer results was undetermined. In other words, better results might not have been obtained using any assessment procedure with the particular sample studied by Derby et al. To examine this possibility, Kahng and Iwata (1999) recently compared outcomes from 50 functional analyses consisting of brief (single) or repeated exposure to test and control conditions. Their results indicated that interpretations based on the brief analyses corresponded with those based on the extended analyses for 33 of the cases (66%).

In spite of the advantages offered by brief functional analyses, including efficiency and objectivity (especially when compared with assessments based solely on verbal report obtained during an interview), there are several limitations. First, single exposures to test conditions may occasionally yield no occurrences of the target behavior, as was observed for 37% of the cases in the Derby et al.

(1992) study. Second, results will be clear only to the extent that individuals can discriminate differences among conditions almost immediately. Finally, processes that may affect behavior over time, such as extinction, may not occur in a single session. For these reasons, repeated exposure to assessment conditions is highly desirable. One way to achieve this goal, while minimizing the time required to conduct an assessment, would consist of running brief but repeated sessions under each condition. The purpose of this study was to determine the extent to which brief exposure to assessment conditions might yield results similar to those obtained from longer exposure by comparing outcomes from functional analyses based on session durations of 15, 10, and 5 min.

METHOD

Participants and Setting

Forty-six adult individuals (23 males, 23 females) who had been referred for assessment and treatment of self-injurious behavior (SIB) or aggression participated. All lived at a state residential facility for persons with developmental disabilities and had been diagnosed with severe or profound mental retardation. These individuals were selected for participation from a pool of clients who had been assessed from 1990 to 1997 based on the following criteria: (a) Assessment consisted of a functional analysis in which 15-min sessions were arranged in a multielement design, (b) the assessment represented the individual's first exposure to a functional analysis, and (c) at least one instance of the target behavior was observed. These data provided the standards for comparison.

All sessions were conducted in individual therapy rooms at a day-treatment program located on the grounds of the residential facility. Therapy rooms contained tables, chairs, and other materials relevant to the various experimental conditions.

Table 1
Response Topographies and Definitions

Response	Definition	Number of participants[a]
SIB		
Body hitting	Forceful (audible) contact of a hand against the body	15
Face scratching	Contact between a fingernail and skin on the face or head	4
Eye poking	Insertion of a finger into the eye socket	1
Hand biting	Contact between the teeth and any part of the hand or fingers	8
Hand mouthing	Insertion of a finger or hand past the plane of the lips	11
Head banging	Forceful (audible) contact between the head and any object	9
Head hitting	Forceful (audible) contact between the hand and head	24
Hair pulling	Closing the fingers or hand around hair strands and pulling	2
Aggression		
Biting	Closure of teeth on skin or clothing of another	3
Hitting	Striking movement against any part of the body of another	7
Scratching	Raking of fingernails on the skin of another	4

[a] Exceeds the total number of participants because some individuals engaged in more than one topography of SIB.

Response Measurement and Interobserver Agreement

Table 1 shows topographies of SIB and aggression (including response definitions) observed throughout the study and the number of participants who engaged in each topography. Trained observers recorded occurrences of these responses on handheld computers, and data were converted to number of responses per minute. Sessions were 15 min in duration and were conducted one to four times daily, usually 5 days per week.

Interobserver agreement was assessed by having a second observer independently record data during 35% of the sessions. Agreement scores were calculated by dividing session time into consecutive 10-s intervals and comparing observers' records. The smaller number of responses was divided by the larger number of responses in each interval; these fractions were summed, divided by the number of intervals in the session, and multiplied by 100%. Mean agreement was 97.1% across participants (range, 83.8% to 100%).

Functional Analysis

All participants were exposed to attention, demand, alone, and play conditions in a multielement design based on procedures described by Iwata et al. (1982/1994). Four individuals were also exposed to a tangible condition based on preliminary information suggesting that problem behavior might be maintained by access to specific leisure items.

Positive reinforcement (attention). An experimenter and participant were present in a room containing leisure materials or food, to which the participant had free access throughout the session. The experimenter ignored the participant throughout the session, except to deliver attention (e.g., "Don't do that, you'll hurt yourself," "Stop, you're hurting me") whenever the participant engaged in a target behavior.

Positive reinforcement (access to tangible items). The participant was provided free access to a preferred leisure item for several minutes prior to the session. When the session began, the experimenter removed the item but kept it in view. When the participant emitted a target behavior, the experimenter delivered the item for 30 s.

Negative reinforcement (escape from demands). The experimenter initiated learning trials every 30 s using a three-prompt pro-

cedure (instruction, followed by demonstration and physical prompting, if necessary). Compliance was followed by praise, whereas occurrence of a target behavior terminated the trial.

Alone. The participant was observed while in a room alone. Leisure materials were unavailable.

Play. An experimenter and participant were present in a room containing leisure materials, to which the participant had free access throughout the session. The experimenter interacted with the participant in a friendly manner at least once every 30 s. Occurrences of target behaviors were ignored.

Data Calculation and Interpretation

The 46 functional analyses based on 15-min sessions were used to generate additional data sets by subtracting (a) the last 5 min and (b) the last 10 min from each session. This resulted in three complete sets of data based on 15-, 10-, and 5-min session durations. Individual graphs were prepared for each functional analysis ($N = 138$), organized in random sequence, and then evaluated individually by six graduate students whose experience in conducting and interpreting functional analysis data ranged from 2 to 5 years. These individuals were blind to participant identity and session duration. When evaluating each graph, the raters selected from attention, tangible, escape, and automatic reinforcement functions (or a combination), or determined that the results were ambiguous. Each of the graphs was then given a designated function based on agreement among at least five of the six raters. We then determined the extent to which functions for the 10- and 5-min data sets matched those obtained for the 15-min data set.

RESULTS

Table 2 summarizes the results obtained from comparisons of the three data sets. A

Table 2
Designated Functions for Data Sets Based on Different Session Durations

Function	Session duration		
	15 min	10 min	5 min
Attention	6	6	6
Tangible	4	4	3
Escape	13	13	12
Automatic	9	9	10
Attention and automatic	1	1	1
Ambiguous	13	13	14
Total	46	46	46

comparison of interpretations based on 10-min and 15-min sessions yielded perfect agreement, whereas interpretations based on 5-min and 15-min sessions yielded three disagreements. Of the three disagreements, two of the 15-min data sets that yielded a clear function (one tangible and one escape) were seen as ambiguous based on 5-min samples, and one of the 15-min data sets that yielded an ambiguous function was identified as indicative of an automatic reinforcement function based on its 5-min sample.

Figure 1 shows an example of one of the data sets in which all three session lengths indicated the same function. Brandy's results show correspondence among interpretations based on 15-, 10-, and 5-min session durations. Her SIB was maintained by attention, and all three graphs show no overlap of data points from the attention condition with any other condition.

Figure 2 shows the three data sets in which there was disagreement between interpretations based on 5-min and 15-min sessions. Results of Eli's functional analysis (left panel) based on data from 15-min and 10-min sessions both showed a clear escape function in that there was no overlap of data points from the demand condition with any other condition. When session duration was reduced to 5 min, however, Eli's SIB decreased noticeably during the first demand session and dropped to zero during the third
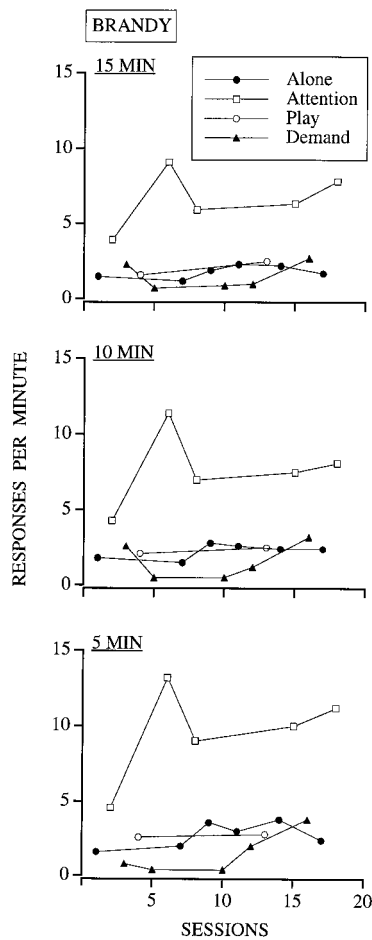
Figure 1.   Example of a functional analysis showing correspondence among data sets based on 15-, 10-, and 5-min session durations (Brandy).

demand session because most (first session) or all (third session) of his SIB occurred during the second half of those sessions. This can be seen in Figure 3 (top panel), which shows a cumulative record of Eli's SIB during the first, second, and third demand sessions. Thus, occurrences of SIB during the first and third sessions were reflected only in the 10-min and the 15-min data sets, yielding an ambiguous outcome for the 5-min data set.

Results of Dan's functional analysis (Figure 2, middle panel) based on 15-min session durations showed SIB occurring only in the tangible condition (three of four tangible

sessions). When session length was reduced to 10 min, SIB disappeared during the second tangible session because all of Dan's SIB occurred during the last 5 min of that session (see Figure 3, middle panel, second session). Nevertheless, noticeable rates of SIB were still evident in two of four tangible sessions of the 10-min sample (Figure 2), yielding a consistent interpretation. By contrast, when session duration was reduced to 5 min, noticeable rates of SIB were evident in only one of the four tangible sessions. The ambiguous outcome based on the 5-min sample was due to the fact that all of Dan's SIB during the third tangible session occurred beyond the 5-min mark (Figure 3, middle panel, third session). Thus, in Dan's case, limiting session duration to 10 min produced a slight and insignificant loss in clarity, whereas limiting session duration to 5 min obscured an originally clear outcome.

The results of Rodney's functional analysis based on 15- and 10-min session durations (Figure 2, right panel) were interpreted as ambiguous due to variable rates of responding in both the demand and alone conditions. When session duration was reduced to 5 min, however, results suggested that Rodney's SIB was maintained by automatic reinforcement. This discrepancy resulted from the fact that almost all of his SIB during the first five demand sessions occurred beyond the 5-min mark (Figure 3, bottom panel) and was therefore undetectable in the functional analysis based on 5-min session durations.

DISCUSSION

We determined the extent to which response patterns observed during functional analyses based on 15-min session durations would remain consistent if session time were reduced to either 10 or 5 min. Examination of 46 sets of data yielded perfect agreement for results based on comparisons of 15- and
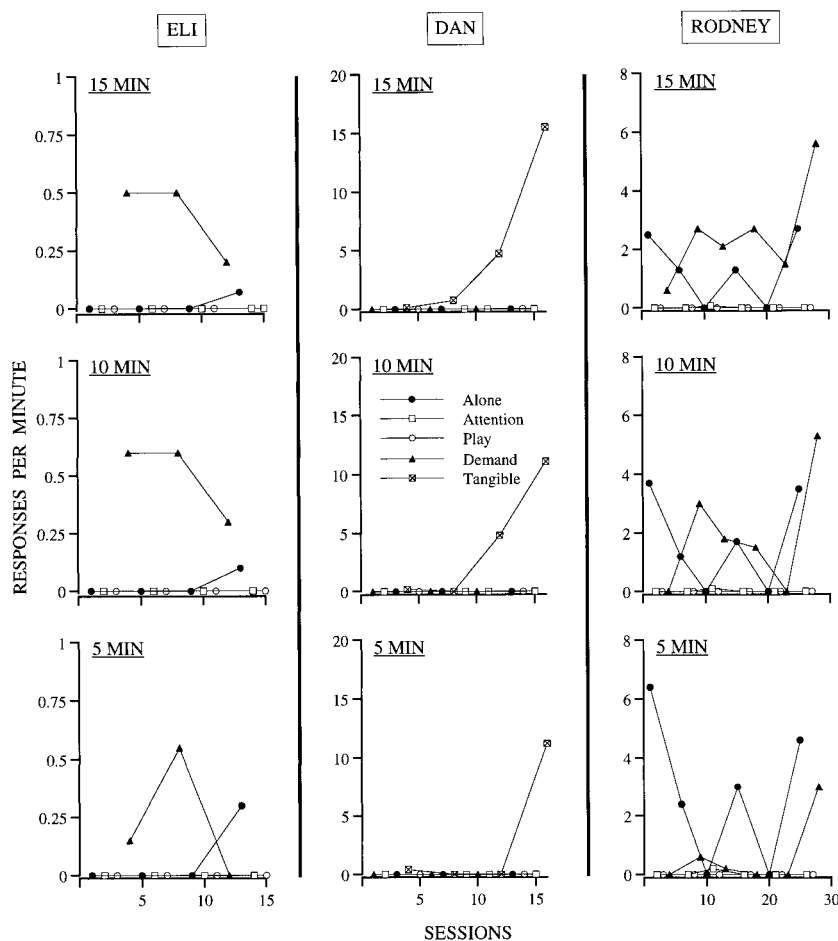
Figure 2.   The three cases showing disagreement between interpretations based on 15- and 5-min session durations.

10-min durations and only three disagreements for results based on comparisons of 15- and 5-min durations. The mean durations (total time) of assessment per individual would have been 6.1 hr, 4.1 hr, and 2.1 hr, respectively, for the 15-, 10-, and 5-min functional analyses. These differences represent noticeable reductions in assessment time with little loss in interpretive power.

Our results suggest that brief repeated exposure to assessment conditions may be sufficient to reveal functional relationships in a large proportion of cases. This conclusion must be tempered, however, because there are certain situations in which brief session durations may be problematic. First, some individuals may not discriminate differences among conditions quickly, especially when different conditions share incidental features. For example, when different conditions are conducted under similar physical arrangements (e.g., same therapist, same location), discrimination may require extensive contact with contingencies, which may be precluded when session duration is brief. Second, extinction during control or nontest conditions (e.g., play, alone) may require extended exposure to the absence of a contingency. Finally, antecedent manipulations that alter establishing operations (EOs) may occasion behavior only after they have been in effect for some time. For example, deprivation
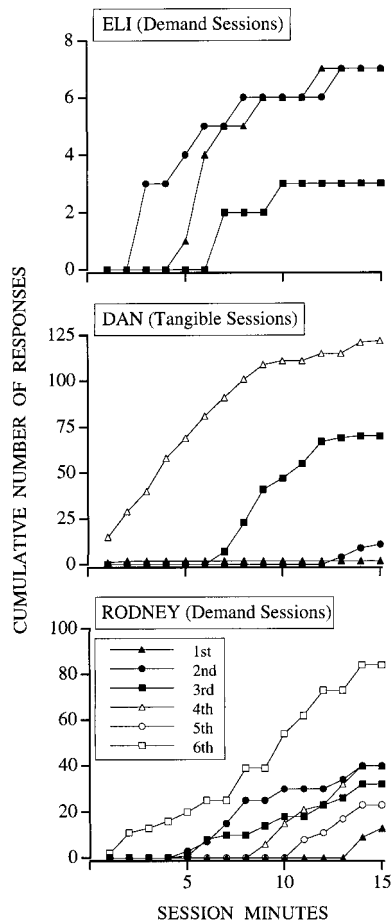
Figure 3. Cumulative records of SIB during Eli's demand sessions (top panel), Dan's tangible sessions (middle panel), and Rodney's demand sessions (bottom panel) based on minute-by-minute response rates.

Rodney's six demand sessions, SIB occurred either never or rarely during the first 5 min, yielding an outcome that was atypical of his behavior under demand conditions of longer duration. One way to detect changes in responding during a session that may reflect discrimination, extinction, and so on, consists of examination of within-session (i.e., minute by minute) response patterns (Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993). However, even this approach would have been inadequate to capture the within-session rate changes in Eli's, Dan's, or Rodney's 5-min data sets because their SIB increased beyond the 5-min mark.

The emergence of SIB toward the latter part of sessions raises a concern about the manner in which comparisons were made in the present study. Although the 15-min data sets reflected the total time of exposure to all conditions, the partial data sets reflected only the first 10 or 5 min of each session, uncorrected for total exposure time. For example, the second data point for a given condition (e.g., demand) in a 5-min data set was based on previous exposure to that condition for 15 instead of 5 min (i.e., the first demand session actually lasted 15 min, but responding during only the first 5 min was reflected in the graph). Thus, an alternative method of comparison could have been based on total exposure time by transforming a single 15-min session into three 5-min sessions. This type of transformation seemed unjustified because, as shown in each of the cumulative graphs in Figure 3 (Rodney's in particular), responding seemed to vary as a function of within-session exposure to a condition rather than total exposure to that condition (i.e., responding was generally low during the first few minutes of a session regardless of how many sessions had been conducted). Even if one were to conduct three independent analyses based on different session durations (5, 10, and 15 min), any effects due to total exposure might influence

from attention or aversive stimulation in the form of task demands may function as a weak EO at the beginning of a session but as a powerful EO as session time increases, or vice versa (Smith, Iwata, Goh, & Shore, 1995).

All of the above factors increase the likelihood that response rates at the beginning of a session may be much different than those at the end of a session, as was observed in each of the three cases for which there was a discrepancy between 5- and 15-min session durations, Rodney's data being the clearest example. In each of the first five of

results of the second and third analyses. Therefore, the approach taken in the current study seemed to represent a reasonable compromise between within-session (local) and within-condition (total) histories of exposure, and revealed the existence of some within-session influences.

The extent to which within-session changes in response rates during functional analyses affect overall session values is unknown, but it has been observed in several studies (Kahng & Iwata, 1999; Vollmer et al., 1993). Given that this problem was observed in only 3 of 46 sets of 5-min data in the present study, one might dismiss it in favor of the increased efficiency afforded by 5-min sessions, which represent 50% and 67% reductions in time compared to session durations of 10 and 15 min, respectively. Alternatively, a more conservative assessment strategy would incorporate session durations of 10 min because interpretations based on these data matched perfectly those based on 15-min durations. It is, of course, possible that durations even longer than 15 min may be required occasionally, and an interesting extension of this study might document the conditions under which extremely long assessment sessions aid in the identification of behavioral function.

REFERENCES

Derby, K. M., Wacker, D. P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25,* 713–721.

Durand, V. M., & Crimmins, D. B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18,* 99–117.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27,* 197–209. (Reprinted from *Analysis and Intervention in Developmental Disabilities, 2,* 3–20, 1982)

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., Rodgers, T. A., Lerman, D. C., Shore, B. A., Mazaleski, J. L., Goh, H., Cowdery, G. E., Kalsher, M. J., McCosh, K. C., & Willis, K. D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *Journal of Applied Behavior Analysis, 27,* 215–240.

Kahng, S., & Iwata, B. A. (1999). Correspondence between outcomes of brief and extended functional analyses. *Journal of Applied Behavior Analysis, 32,* 149–159.

Lennox, D. B., & Miltenberger, R. G. (1989). Conducting a functional assessment of problem behavior in applied settings. *Journal of the Association of Persons with Severe Handicaps, 14,* 304–311.

Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., & DeRaad, A. (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. *Journal of Applied Behavior Analysis, 24,* 509–522.

Pyles, D. A. M., Riordan, M. M., & Bailey, J. S. (1997). The stereotypy analysis: An instrument for examining environmental variables associated with differential rates of stereotypic behavior. *Research in Developmental Disabilities, 18,* 11–38.

Repp, A. C., Singh, N. N., Olinger, E., & Olson, D. R. (1990). The use of functional analyses to test causes of self-injurious behavior: Rationale, current status, and future directions. *Journal of Mental Deficiency Research, 34,* 95–105.

Smith, R. G., Iwata, B. A., Goh, H., & Shore, B. A. (1995). Analysis of establishing operations for self-injury maintained by escape. *Journal of Applied Behavior Analysis, 28,* 515–535.

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). Within-session patterns of self-injury as indicators of behavioral function. *Research in Developmental Disabilities, 14,* 479–492.

STUDY QUESTIONS

1. What are some strengths and limitations associated with functional analyses that consist of single exposures to assessment conditions (i.e., brief functional analyses)?

2. How were data for the "different" functional analyses generated and interpreted?

3. To what extent were interpretations based on 10- and 5-min data sets consistent with those based on the 15-min data sets?

4. In the results section, what characteristics of the data are specifically described as influential in interpretation?

5. What was the source of discrepancy for all of the obtained disagreements?

6. The authors mention several conditions under which brief session durations may be problematic. What are these conditions, and which ones may have influenced the present results?

7. Based on the results of this study, what was the authors' recommendation regarding optimal session duration when conducting typical functional analyses?

8. The above recommendation must be tempered by the fact that data, as presented in this study, were not truly representative of behavior that might have been observed had assessment sessions been limited to 10- or 5-min durations. Why was this the case?

Questions prepared by Gregory Hanley and Rachel Thompson, The University of Florida