According to recent studies, the percentage of U.S. adults who are functionally illiterate—here defined as the inability to read or write in any practical capacity, such as to read a job application—as recently as 2013 was as high as 14% (Literacy Project Foundation). This figure includes 24% of the adult U.S. black population, 41% of the Hispanic population, and a staggering 70% of the prison population (U.S. Department of Education, 2015). Naturally, the challenges incurred by poor English literacy in the United States spread over a range of circumstances, making it difficult for illiterate adults to find employment, get an education, obtain medical care, or navigate many other aspects of day-to-day life. These difficulties extend to students as well, with as many as 64% of U.S. eighth graders unable to read at their grade level on standardized testing (National Center for Education Statistics, 2013).

The long-term objective of this research proposal is to create educational tools that can improve education in adult literacy and other academic fields. To this end, this proposal puts forth two main goal products: an educational system called reMind, and an add-on passage-comprehension module called CAPITAL Passages.

Primarily, reMind will be a system that allows instructors to deliver practice and evaluation questions to students and students to practice study materials at their own pace. The primary algorithmic distinction of this system, which will be studied and implemented by an external team, is the spaced repetition of study materials at calculated intervals to promote long-term learning as well as learning material for a specific deadline. The portion of this project implemented by this team will involve the development of a user interface including a website for instructors to collaborate with students and other instructors to amass a large collection of academic questions over many academic disciplines, and a mobile application for students to study their materials in a convenient way and receive immediate feedback about the quality of their responses.

More connected to the goal of improving literacy is the CAPITAL Passages system, which will incorporate research into natural language processing to generate coherent questions that test reading comprehension on the sentence and multiple-sentence semantic level, as opposed to on the simpler word-recognition level. Passages will be an add-on module to the user interface of reMind, meaning that the algorithms it uses to generate its questions will also produce the correct answers to those questions, as well as incorrect distractor answers when appropriate, in order for students’ responses to be evaluated immediately.

The **intellectual merit** of the project is the linguistic and algorithmic challenge of creating the Passages system. Passages will take the current abilities of natural language processing further by moving on to comprehension questions incorporating multiple sentences within a paragraph, while retaining knowledge of the correct answer even when question creation becomes complex. In addition, Passages will combine different types of question generation—such as an algorithm for true/false questions and a separate algorithm for multiple choice questions—into the same tool in order to generate more and higher-quality questions from the same passage. This system will allow Passages to test reading comprehension in a deeper and more rigorous way than a system that simply asks factual questions about one sentence. Potential benefits of this project will be in the fields of natural language processing and education.

The **broader impact** of the project lies in its potential to improve adult literacy and existing education schemes. Through automatic generation, the project will provide a low-maintenance system which can test and train whole-passage comprehension. The project will also allow low-literacy individuals to build literacy-related life skills, which are often a barrier to pursuing education. These improved skills will lead to increased ability of these groups (often impoverished, ethnic minority, or immigrant populations) to participate in all areas of education and industry. Furthermore, the chief value of the reMind system will be its capacity to amass, generate (through modules such as Passages), and store databases of academic questions to be shared among instructors. While the development of these systems will be specific to the English language, some attention to linguistic differences would allow the algorithms to be adapted for other languages.